Reconstruction of Molecular Orbitals from Photoemission Data with Iterative Phase Retrieval Algorithms

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Pavlo Kliuiev
aus der Ukraine

Promotionskommission

Prof. Dr. Jürg Osterwalder (Vorsitz)
Dr. Luca Castiglioni (Leitung der Dissertation)
Dr. Matthias Hengsberger
Prof. Dr. Hans-Werner Fink
PD Dr. Achim Schöll

Zürich, 2018

To my parents.

ABSTRACT

The present work addresses the reconstruction of molecular orbitals from angle-resolved photoelectron spectroscopy (ARPES) data by means of orbital tomography. Provided ARPES data are recorded from an ordered layer of molecules consisting of light atoms only, final state scattering effects can be neglected and the ARPES data can be related to the squared modulus of the Fourier transform of the initial state wave function. The wave function can be retrieved by the inverse Fourier transform, if the phase distribution in the detector plane is known. We suggest finding the phase distribution in a robust manner by exploiting an analogy between the phase problem in coherent diffraction imaging (CDI) and orbital tomography. We performed an optical CDI experiment and used the phase retrieval algorithms of Gerchberg-Saxton, Fienup and Marchesini to reconstruct micrometer-sized structures solely from their diffraction patterns. By applying the same algorithms to ARPES data, we reconstructed an orbital of pentacene on Ag(110). A further point that we address is the issue of the background subtraction, for which we propose the use of a signal restoration routine based on the maximisation of mutual information algorithm. In addition, we solved the problem of finding the geometrical centre in reconstructed orbital distributions by modifying the phase retrieval procedure. By comparison with photoelectron holography, we show that a reconstructed two-dimensional orbital distribution can be interpreted as a superposition of the "in-focus" orbital distribution evaluated at the $z=0$ plane and "out-of-focus" distributions evaluated at other $z = \text{const}$ planes. Then, we extend the orbital tomography to cobalt pyrphyrin and pyrphyrin molecules on Ag(110). By comparing the experimental ARPES data with the density functional theory simulations, we find that the molecules are ordered in multiple rotational domains on the surface of the Ag(110) crystal. This issue challenges the reconstruction of individual orbital distributions from the experimental data. Finally, we performed first time-resolved experiments on a multilayer of pentacene on Ag(110), in which the ultimate goal is the observation of excited state dynamics and subsequent reconstruction of orbital distributions at different moments in time.

ZUSAMMENFASSUNG

Die vorliegende Arbeit befasst sich mit der Rekonstruktion von Molekülorbitalen aus den experimentellen Daten der winkelaufgelösten Photoemission (engl. *angle-resolved photoelectron spectroscopy*, ARPES) mittels der sogenannten Orbitaltomographie. Das Verfahren funktioniert unter der Voraussetzung, dass die experimentellen Daten von einer geordneten Schicht der organischen Moleküle akquiriert werden, so dass die Endzustand-Streueffekte vernachlässigt werden können. In diesem Fall entsprechen die ARPES Daten der quadrierten Amplitude der Fourier Transformation der molekularen Wellenfunktion im Anfangszustand. Dieser Zusammenhang ermöglicht die Rekonstruktion der Wellenfunktion, falls man die Phasenverteilung in der Ebene des Detektors kennt und die inverse Fourier Transform der komplexen Signalverteilung in der Detektorebene berechnet. Wir schlagen eine Methode für die Lösung des Phasenproblems vor, die auf der Analogie zum Phasenproblem in *Coherent Diffraction Imaging* basiert. Wir führten ein Experiment aus, in dem die Amplituden- und Phasenverteilung der mikrometergrossen Strukturen durch die *phase retrieval* Algorithmen von Gerchberg-Saxton, Fienup und Marchesini rekonstruiert wurden. Wir benützten dieselben Algorithmen für die erfolgreiche Rekonstruktion eines Orbitals von Pentacen auf Ag(110). Ein weiterer Punkt, den wir ansprechen, ist das Problem der Hintergrundsubtraktion in ARPES Daten. Wir schlagen eine Prozedur vor, in der das Signal und der Hintergrund mittels eines *maximisation of mutual information* Algorithmus in räumliche Korrespondenz gebracht werden, um eine anschliessende Subtraktion durchzuführen. Wir verbesserten das *phase retrieval*, in dem die Mitte der rekonstruierten Verteilungen eindeutiger bestimmt wird. Im Vergleich mit Photoelektronenholographie stellen wir fest, dass die zwei-dimensionalen rekonstruierten Molekülorbitale eine Superposition der *im-Fokus* und *aus-dem-Fokus* Komponenten zu interpretieren sind. Ausserdem wird die Orbitaltomographie zu Cobalt-Pyrphyrin und Pyrphyrin Moleküle auf Ag(110) erweitert. Eine umfangreiche Analyse dieses Systems zeigt, dass sich Cobalt Pyrphyrin und Pyrphyrin Moleküle auf der Oberfläche des Ag(110) Kristalls in drei, bzw. zwei Domänen ordnen. Schlussendlich zeigen wir die ersten zeitaufgelösten Experimente, in denen eine Doppellage von Pentacen auf Ag(110) erforscht wird, wobei das Endziel die zeitaufgelöste Rekonstruktion der Molekülorbitale ist.

PUBLICATIONS

PUBLICATIONS RELATED TO THIS DISSERTATION

[PK1] **Orbital tomography of complex macrocycles**. P. Kliuiev et al. *In preparation* (2018)

[PK2] **Algorithms and image formation in orbital tomography**. P. Kliuiev, T. Latychevskaia, C. Metzger, M. Grimm, G. Zamborlini, M. Jugovac, A. Schöll, J. Osterwalder, M. Hengsberger, L. Castiglioni. *Submitted* (2018)

[PK3] **Application of iterative phase-retrieval algorithms to ARPES orbital tomography**. P. Kliuiev, T. Latychevskaia, J. Osterwalder, M. Hengsberger, L. Castiglioni. *New J. Phys.* **18** 093041 (2016)

OTHER PUBLICATIONS

[PK4] **Effective mass effect in attosecond electron transport**. L. Kasmi, M. Lucchini, L. Castiglioni, P. Kliuiev, J. Osterwalder, M. Hengsberger, L. Gallmann, P. Krueger and U. Keller. *Optica* **4** (12) 1492–1497 (2017)

[PK5] **Light-matter interaction at surfaces in the spatiotemporal limit of macroscopic models**. M.Lucchini, L. Castiglioni, L. Kasmi, P.Kliuiev, A. Ludwig, M. Greif, J. Osterwalder, M. Hengsberger, L. Gallmann, U. Keller. *Phys. Rev. Lett.* **115** (13) 137401 (2015)

CONTENTS

I THEORETICAL BACKGROUND, METHODS, MATERIALS AND EQUIPMENT

II RESULTS

LIST OF ABBREVIATIONS

2PPE	Two photon photoemission
2D	Two-dimensional
3D	Three-dimensional
AC	Analysis chamber
AG	Adsorption geometry
ARPES	Angle-resolved photoelectron spectroscopy
CoPyr	Cobalt Pyrphyrin
CCD	Charge-coupled device
CDI	Coherent diffraction imaging
CPA	Chirped pulse amplification
CW	Continuous wave
DFT	Density functional theory
EDC	Energy distribution curve
ER	Error reduction (algorithm)
FFT	Fast Fourier transform
FWHM	Full width at half maximum
GaAs	Gallium arsenide
GDD	Group delay dispersion
GVD	Group velocity dispersion
HHG	High harmonic generation
HOMO	Highest occupied molecular orbital
IAC	Independent-atomic-centre (approximation)
IR	Infrared
LEED	Low energy electron diffraction
LUMO	Lowest unoccupied molecular orbital
MCP	Micro-channel plate
MI	Mutial information
ML	Monolayer
OPA	Optical parametric amplifier
PAD	Photoelectron angular distribution
PDOS	Partial density of states
PEEM	Photoemission electron microscope
PC	Preparation chamber
PC-HIO	Phase-constrained hybrid input-output (algorithm)
Pyr	Pyrphyrin
PW	Plane wave
RegA	Regenerative amplifier
RMS	Root mean square

SEM	Scanning electron microscope
SHG	Second harmonic generation
STM	Scanning tunnelling microscope
VBM	Valence band maximum
XPD	X-ray photoelectron diffraction
XPS	X-ray photoelectron spectroscopy
XUV	Extreme ultraviolet
UPS	Ultraviolet photoelectron spectroscopy
UZH	University of Zurich
UV	Ultraviolet

LIST OF COMMON SYMBOLS AND PHYSICAL CONSTANTS

$\boldsymbol{A}$	Vector potential of electromagnetic radiation
$c \approx 3 \cdot 10^8$ m/s	Speed of light in vacuum
$e \approx 1.602 \cdot 10^{-19}$ C	Elementary charge
E_b	Electron binding energy
E_f	Electron final state energy
E_F	Fermi energy
E_i	Electron initial state energy
E_kin	Electron kinetic energy
E_R	Root-mean-square error metric in reciprocal space
$\mathcal{F}$	Fourier transform
$\mathcal{F}^{-1}$	Inverse Fourier transform
$h\nu$, $\hbar\omega$	Photon energy
$\hbar \approx 1.054 \cdot 10^{-34}$ J · s	Reduced Planck's constant
$\hat{\mathcal{H}}$	Time-dependent Hamiltonian
H'_{mk}	Matrix element of the perturbation operator
$\boldsymbol{k}$	Photoelectron wave vector in vacuum
$\boldsymbol{k}_s$	Photoelectron wave vector in solid
$\boldsymbol{k}_\parallel$	Parallel component of photoelectron wave vector
$\boldsymbol{k}_\perp$	Perpendicular component of photoelectron wave vector
$\boldsymbol{K}$	Wave vector of light
$m \approx 9.109 \cdot 10^{-31}$ kg	Electron rest mass
$\hat{\boldsymbol{p}}$	Electron momentum operator
$\boldsymbol{r}$	Radius vector
V_0	Mean inner potential
λ	Wavelength of light in vacuum
λ_dB	De Broglie wavelength of electrons
ν, $\omega = 2\pi\nu$	Frequency and angular frequency of light
Ø	Linear oversampling ratio
σ	Oversampling ratio
σ_0	Standard deviation of Gaussian distribution
θ	Polar detection angle in vacuum
θ_s	Polar detection angle in solid

ϕ	Phase distribution in real space
Φ	Phase distribution in reciprocal space
Φ_0	Work function
$\psi(x, y, z)$	Wave function distribution in real space
$\Psi(k_x, k_y, k_z)$	Wave function distribution in reciprocal space
$\lvert\psi_\mathrm{f}\rangle$	Final state wave function
$\lvert\psi_\mathrm{i}\rangle$	Initial state wave function

INTRODUCTION AND MOTIVATION

Spatiotemporal visualisation of electron dynamics in molecules is necessary for further advancement in efficiency of organic light-emitting diodes, photovoltaic cells and photocatalytic devices. Angle-resolved photoelectron spectroscopy (ARPES) of well-ordered layers of organic large planar molecules on single crystal substrates provides means for reconstruction of initial state molecular wave functions [1–5]. Within the plane wave (PW) approximation for the photoelectron final state [1, 2, 6], the ARPES intensity distribution is proportional to the squared modulus of the Fourier transform of the initial state wave function and the latter can therefore be retrieved via inverse Fourier transform of the square root of the ARPES data, provided the phase of the photoelectron wave distribution in the detector plane is known. The phase distribution may be gained from the parity of the wave function [1], dichroism measurements [7] or iteratively, employing knowledge about the shape of the wave function [3]. Up to now, this technique has been successfully applied for reconstruction of orbital distributions of polycyclic aromatic hydrocarbons and organic dye molecules [1, 3, 5, 7] and become known as orbital tomography. In this work, we advance this technique both methodologically by making improvements in orbital reconstruction routines and experimentally by extending it to a broader class of organic molecules and performing pump-probe measurements to characterise temporal behaviour of excited molecular states.

As the phase distribution in the detector plane is always lost in the measurement, this issue constitutes a fundamental problem in orbital tomography, called the phase problem. We suggest solving the phase problem in a robust manner by employing its analogy to the phase problem in coherent diffraction imaging (CDI) [8]. Given the far field optical or photoelectron intensity distribution fulfils the oversampling condition [9, 10], both the amplitude and the phase of the object distribution can be reconstructed from the modulus of its Fourier transform without any information about shape and symmetry properties of the object by applying state-of-the-art phase retrieval algorithms used in CDI [11–13]. This analogy is outlined in part I of the dissertation, where the theoretical fundamentals of photoemission (chapter 1), light diffraction (chapter 2) as well as a detailed description of the iterative phase retrieval algorithms (chapter 3) are provided. Methods and equipment used in the current work are presented in chapter 4.

Part II contains the results. In chapter 5, we demonstrate the results of a CDI experiment performed on micrometer-sized structures [5]. This experiment was designed as an optical analogue in order to validate the phase retrieval algorithms used in CDI. We reconstructed the amplitude and phase distributions of the micrometer-sized structures from their far field diffraction patterns using the shrinkwrap algorithm [12] and a combination of the phase-constrained hybrid input-output (PC-HIO) and error reduction (ER) algorithms [11, 13]. By applying the same reconstruction procedure to a set of ARPES data recorded from a sub-monolayer of pentacene molecules on the Ag(110) substrate, we reconstructed both the amplitude and the phase of the lowest unoccupied molecular orbital (LUMO) of pentacene [5] solely from the raw photoemission data without any symmetrisation or assumptions about symmetry or shape of the orbital distribution.

In many cases, a simple treatment of the experimental data does not suffice. For example, the presence of substrate background features that do not belong to molecular orbital distribution need substantial amount of image processing, prior to application of the phase retrieval algorithms. In chapter 6, we address the issue of substrate background subtraction by suggesting a signal restoration procedure based on the maximisation of mutual information [14]. We also address the problem of finding the geometrical centre in reconstructed orbital distributions by proposing improvements to our phase retrieval routine. We demonstrate the feasibility of the suggested procedures on a set of ARPES data recorded from the highest occupied molecular orbital (HOMO) of pentacene on the Ag(110) substrate. We compare the iterative reconstructions with the results of density functional theory (DFT) [16, 17] simulation by modifying the DFT data in accordance with the experimental conditions. Most importantly, by making a comparison with photoelectron holography [18], we interpret two-dimensional (2D) reconstructed orbital distributions as a superposition of “in-focus” and “out-of-focus” contributions from the original three-dimensional (3D) orbitals. This remarkable result means that the full 3D orbital distributions can be reconstructed solely from a single set of 2D ARPES data, provided the axial resolution is high enough.

In chapter 7, we extend orbital tomography to molecular systems relevant to catalysis: pyrphyrin (Pyr) and cobalt-pyrphyrin (CoPyr) molecules on Ag(110). Pyr molecule is used as a carrier for catalytically active atoms of transition metals and its metalated form CoPyr is a promising water reduction catalyst in photocatalytic water splitting [19]. Ability to reconstruct molecular orbitals in these systems is therefore important for validation of theoretical models [20] and extension of orbital tomography to a broader class of molecules. Photoelectron angular distribution (PAD) maps of Pyr and CoPyr on Ag(110) were recorded using the photoemission electron microscope (PEEM) at the NanoESCA beamline at the Elettra synchrotron radiation facility in Trieste, Italy. This beamline provides photon energies in the range of 25-35 eV and higher, which is essential for the reconstruction of orbital distribution from photoemission data, as only at these photon energies, corresponding to photoelectron momenta of $\pm$ 2...3 Å^{-1}, most of the molecular features in PADs become present. In Pyr and CoPyr data, we identify several molecular states between the Fermi level and the onset of d-bands. The characterisation of the samples with low energy electron diffraction (LEED), predicted the existence of two molecular lattices on the Ag(110) in both types of molecular systems and DFT simulations predicted the existence of three adsorption geometries. To exactly determine the spatial orientation of the individual molecules, we performed x-ray photoelectron diffraction (XPD) measurements of Pyr and CoPyr on N 1s and Co 2p states. However, insufficient signal-to-noise ratio and issues with the Ag(110) crystal challenged unambiguous determination of the molecular orientation. Thus, to determine the orientation of individual molecules, we used the results of LEED and simulations done using DFT in the frameworks of PW and independent-atomic-centre (IAC) [6] approximations. By comparing the experimental ARPES data with the simulated PADs, and using the information about possible spatial arrangements of the molecules obtained from LEED, we found that CoPyr and Pyr molecules most possible order themselves on the surface of Ag(110) crystal in three and two rotational domains, respectively. This challenges the reconstruction of orbital distributions from the ARPES data and highlights the importance of a single orientation of molecules if one aims at reconstruction of molecular orbitals with phase retrieval algorithms.

Finally, chapter 8 presents the first attempt of time-resolved ARPES measurements using visible pump and extreme ultraviolet (XUV) probe laser pulses. As a proof-of-principle, we show the energy vs. momentum photoelectron distributions recorded from a multilayer of pentacene molecules on the Ag(110) substrate prior and at the moment of excitation, thus, making the first step towards time-resolved orbital tomography.

Part I

THEORETICAL BACKGROUND, METHODS, MATERIALS AND EQUIPMENT

1

PHOTOEMISSION

> *"Could anything at first sight seem more impractical than a body which is so small that its mass is an insignificant fraction of the mass of an atom of hydrogen? –which itself is so small that a crowd of these atoms equal in number to the population of the whole world would be too small to have been detected by any means then known to science."*
>
> Sir Joseph John Thomson, 1934

1.1 INTRODUCTION

It was in 1887, when Heinrich Hertz, working on generation and detection of electromagnetic waves, noticed that a spark in a gap between two pieces of metal is more readily observed under the illumination by ultraviolet light [21]. The electron was discovered 10 years later by J. J. Thomson and within the next two years he demonstrated that the negatively charged particles emitted in the experiment of Hertz were electrons [22]. In 1905, Albert Einstein suggested that light consists of quanta – photons – that transfer their energy to electrons and thereby evoke them from a solid [23]. The kinetic energy of a photoemitted electron in vacuum, E_{kin}, equals to

$$E_{\mathrm{kin}} = h\nu - E_{\mathrm{b}} - \Phi_0, \tag{1.1}$$

where $h\nu$ is the energy of a photon, E_{b} is the electron binding energy referenced with respect to the Fermi energy, E_{F}, where it is equal to zero, and Φ_0 is the work function of the sample. In 1964, Berglund and Spicer [24] introduced the *three-step model* of photoemission schematically sketched in Fig. 1.1(A). This model employs a simple picture of non-interacting electrons and divides the photoemission process into three steps [25, 26].

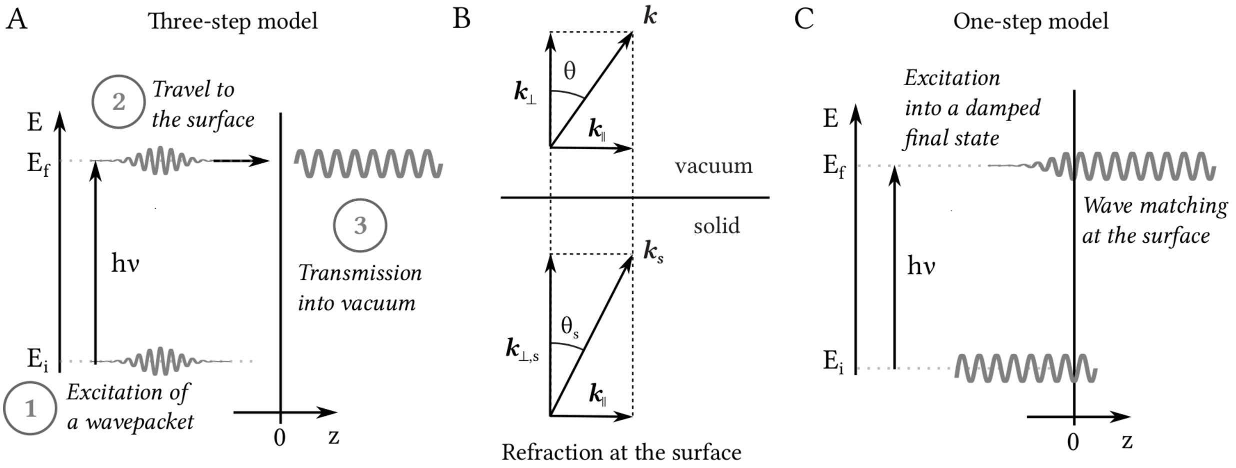


Figure 1.1: (A) – Three-step model of photoemission. E_{i} and E_{f} are initial and final state energies, respectively. $h\nu$ is the photon energy. (B) – Electron refraction at surface. $\boldsymbol{k}_s$ is the photoelectron wave vector in solid, $\boldsymbol{k}_{\perp,s}$ is the perpendicular component of the photoelectron wave vector in solid, $\boldsymbol{k}_{\parallel}$ is the parallel component of the photoelectron wave vector, $\boldsymbol{k}_{\perp}$ is the perpendicular component of the photoelectron wave vector in vacuum. θ is the detection angle in vacuum, and θ_s the corresponding angle in solid. (C) – One-step model of photoemission.

1. Upon absorption of photons with energy $h\nu$, electrons are excited from an initial state $|\psi_{\mathrm{i}}\rangle$ with the energy $E_{\mathrm{i}} = E_{\mathrm{b}}$ to a final state $|\psi_{\mathrm{f}}\rangle$ with the energy $E_{\mathrm{f}} = E_{\mathrm{kin,s}} = h\nu - E_{\mathrm{b}}$ and the momentum $\hbar|\boldsymbol{k}_{\mathrm{s}}| = \sqrt{2mE_{\mathrm{kin,s}}}$. Here E_{b} is the binding energy and $E_{\mathrm{kin,s}}$ is the electron kinetic energy in solid.
2. Electrons travel through the solid towards the surface, enduring elastic and inelastic collisions.
3. Finally, electrons escape into vacuum, where they are detected.

Due to the mean inner potential, V_0, inside the solid, the electrons are refracted at the surface, so that the component of the photoelectron wave vector parallel to the surface, $\boldsymbol{k}_{||}$, is conserved, and the component of the wave vector normal to the surface, $\boldsymbol{k}_{\perp,s}$, is reduced, as it is shown in Fig. 1.1(B). Due to conservation of $\boldsymbol{k}_{||}$, this leads to a refraction law analogous to Snell's law in optics:

$$\frac{\sin\theta_s}{\sin\theta} = \sqrt{\frac{E_{\text{kin}}}{E_{\text{kin,s}}}} = \sqrt{\frac{E_{\text{kin}}}{E_{\text{kin}} + V_0}}, \tag{1.2}$$

where E_{kin} is the electron energy in vacuum, θ is the detection angle with respect to the surface normal in vacuum and θ_s is the corresponding angle in the solid. The parallel component of the photoelectron momentum can be computed as $\hbar|\boldsymbol{k}_{||}| = \sqrt{2mE_{\text{kin}}}\sin\theta$. The perpendicular component in vacuum is then $\hbar|\boldsymbol{k}_\perp| = \sqrt{2mE_{\text{kin}}}\cos\theta$. Estimation of the perpendicular component in the solid requires knowledge about the inner potential though: $\hbar|\boldsymbol{k}_{\perp,s}| = \sqrt{2m(E_{\text{kin}} + V_0)}\cos\theta_s$.

The more accurate quantum mechanical description of photoemission is done using the *one-step model*, in which electrons are excited from an initial state $|\psi_{\text{i}}\rangle$ into a damped final state $|\psi_{\text{f}}\rangle$ near the surface [27], as sketched in Fig. 1.1(C). The final state of the photoelectron is described as the time-reversal of a free electron impinging on the surface from vacuum, as in a low energy electron diffraction (LEED), and is thereby called time-reversed LEED state. To account for the inelastic scattering of electrons in the solid, a propagating free electron state in vacuum is matched to the damped final state in solid [27].

1.2 QUANTUM-MECHANICAL DESCRIPTION OF PHOTOEMISSION

From a quantum mechanical point of view, the rate at which the transition from the initial state $|\psi_{\text{i}}\rangle$ into the final state $|\psi_{\text{f}}\rangle$ takes place is given by Fermi's Golden Rule. Typically it is derived in the framework of time-dependent first order perturbation theory[1]. One begins by considering a time-dependent Hamiltonian $\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{\mathcal{H}}'(t)$, where $\hat{\mathcal{H}}_0$ is the Hamiltonian of an unperturbed system and $\hat{\mathcal{H}}'(t)$ is a time-dependent perturbation. We look for solutions of the perturbed Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial\Psi(\boldsymbol{r},t)}{\partial t} = (\hat{\mathcal{H}}_0 + \hat{\mathcal{H}}'(t))\Psi(\boldsymbol{r},t) \tag{1.3}$$

of the form

$$\Psi(\boldsymbol{r},t) = \sum_k a_k(t)\Psi_k^{(0)}(\boldsymbol{r},t), \tag{1.4}$$

where the coefficients $a_k(t)$ are the functions of time only and the wave functions $\Psi_k^{(0)}(\boldsymbol{r},t)$ are the solutions of the unperturbed Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial\Psi_k^{(0)}(\boldsymbol{r},t)}{\partial t} = \hat{\mathcal{H}}_0\Psi_k^{(0)}(\boldsymbol{r},t). \tag{1.5}$$

By substituting Eq.1.4 into Eq.1.3 and using Eq. 1.5, we obtain

$$\mathrm{i}\hbar\sum_k \Psi_k^{(0)}(\boldsymbol{r},t)\frac{da_k(t)}{dt} = \sum_k a_k(t)\hat{\mathcal{H}}'(t)\Psi_k^{(0)}(\boldsymbol{r},t). \tag{1.6}$$

By multiplying both sides with $\Psi_m^{(0)*}(\boldsymbol{r},t)$ from the left hand side[2], integrating over the whole configuration space and taking into account the orthogonality of the wave functions of the unperturbed system $\Psi_k^{(0)}(\boldsymbol{r},t)$, we obtain

$$\mathrm{i}\hbar\frac{da_m(t)}{dt} = \sum_k H'_{mk}(t)a_k(t), \tag{1.7}$$

where

$$H'_{mk}(t) = \int \Psi_m^{(0)*}(\boldsymbol{r},t)\hat{\mathcal{H}}'(t)\Psi_k^{(0)}(\boldsymbol{r},t)d\xi \tag{1.8}$$

---

[1] We follow the derivation in the books of Landau and Lifshitz [28] and Levich [29].

[2] The asterisk denotes complex conjugation.

are the matrix elements of the perturbation operator and ξ denotes the coordinate of the configuration space. Eq. 1.7 is exact and equivalent to Eq. 1.3, since the coefficients $a_k(t)$ fully determine the wave function $\Psi(\boldsymbol{r},t)$.

In order to simplify the solution of Eq. 1.7, we make use of the fact that the perturbation is weak. Let us assume that the unperturbed wave function was in the n^{th} stationary state. Then at $t \leq 0$ the coefficients in Eq. 1.4 are $a_k(0) = \delta_{kn}$, where δ_{kn} is the Kronecker delta. We look for the coefficients $a_k(t)$ at $t > 0$ of the form

$$a_k(t) = a_k^{(0)}(t) + a_k^{(1)}(t), \tag{1.9}$$

where $a_k^{(0)}(t) = a_k(0) = \delta_{kn}$ and the higher order coefficients were neglected. Upon substitution of Eq. 1.9 into Eq. 1.7, we obtain

$$\mathbb{i}\hbar\frac{da_k^{(1)}(t)}{dt} = \sum_k H'_{mk}(t)a_k^0(t) = H'_{mn}(t). \tag{1.10}$$

By integrating Eq. 1.10, we obtain

$$a_m^{(1)}(t) = -\frac{\mathbb{i}}{\hbar}\int_0^t H'_{mn}(t)dt. \tag{1.11}$$

Because at $t = 0$ the system was in the n^{th} stationary state, $|a_m^{(1)}(t)|^2$ determines the probability of transition from the n^{th} into the m^{th} state during time t: $W_{mn}(t) = |a_m^{(1)}(t)|^2 \equiv |a_{mn}^{(1)}(t)|^2$, where the second index n highlights the initial state of the system.

We are interested in quasi-resonant transitions under the action of a periodic perturbation from the initial state i with the energy $E_{\mathrm{i}}^{(0)}$ to the final state f. We assume the perturbation to be of the form $H'_{\mathrm{fi}}(t) = M_{\mathrm{fi}}e^{\mathbb{i}(\omega_{\mathrm{fi}}-\omega)t}$ and the energy of the final state E_{f} to be very close to the resonance energy $E_{\mathrm{i}}^{(0)} + \hbar\omega$, where $\omega_{\mathrm{fi}} = (E_{\mathrm{f}} - E_{\mathrm{i}})/\hbar$ is the angular frequency. Under these assumptions, the integral in Eq. 1.11 becomes

$$a_{\mathrm{fi}}^{(1)}(t) = -\frac{\mathbb{i}}{\hbar}\cdot\int_0^t H'_{\mathrm{fi}}(t)dt = -M_{\mathrm{fi}}\frac{e^{\mathbb{i}(\omega_{\mathrm{fi}}-\omega)t} - 1}{\hbar(\omega_{\mathrm{fi}} - \omega)} \tag{1.12}$$

The transition probability $W_{\mathrm{fi}}(t)$ then equates to

$$W_{\mathrm{fi}}(t) = \left|a_{\mathrm{fi}}^{(1)}\right|^2 = |M_{\mathrm{fi}}|^2\,\frac{4\sin^2\left(\frac{\omega_{\mathrm{fi}}-\omega}{2}t\right)}{\hbar^2(\omega_{\mathrm{fi}} - \omega)^2} \tag{1.13}$$

or, in the limit of large t, to

$$\lim_{t\to\infty} W_{\mathrm{fi}}(t) = \frac{1}{\hbar^2}\,|M_{\mathrm{fi}}|^2\,\pi t\delta\left(\frac{\omega_{\mathrm{fi}} - \omega}{2}\right) = \frac{2\pi}{\hbar}\,|M_{\mathrm{fi}}|^2\,\delta\left(E_{\mathrm{f}} - E_{\mathrm{i}}^{(0)} - \hbar\omega\right)t, \tag{1.14}$$

where the term with the delta function[3] highlights the conservation of energy. The transition probability in a unit of time is then given by

$$w_{\mathrm{fi}} = \frac{2\pi}{\hbar}\,|M_{\mathrm{fi}}|^2\,\delta(E_{\mathrm{f}} - E_{\mathrm{f}}^{(0)} - \hbar\omega) = \frac{2\pi}{\hbar}\,\left|\langle\psi_{\mathrm{i}}\,|\,\hat{H}'\,|\,\psi_{\mathrm{f}}\rangle\right|^2\,\delta(E_{\mathrm{f}} - E_{\mathrm{i}}^{(0)} - \hbar\omega), \tag{1.15}$$

where the transition matrix element M_{fi} is written explicitly via the initial and final state wave functions. $\hat{H}'$ is the non-oscillatory part of the perturbation Hamiltonian $\hat{\mathcal{H}}'$ and can be found from the full Hamiltonian $\hat{\mathcal{H}}$ by considering the interaction of an electron with an electromagnetic field. For this, we write the unperturbed part of the full Hamiltonian as

$$\hat{\mathcal{H}}_0 = \frac{\hat{\boldsymbol{p}}^2}{2m} + V(\boldsymbol{r}), \tag{1.16}$$

---

[3] $\lim_{t\to\infty}\frac{\sin^2\alpha t}{\pi t\alpha^2} = \delta(\alpha)$, $\delta(cx) = \delta(x)/c$

where $\hat{\boldsymbol{p}} = \frac{\hbar}{\mathbb{i}}\nabla$ is the electron momentum operator and $V(\boldsymbol{r})$ is a periodic potential. By replacing the momentum operator of the electron by $\hat{\boldsymbol{p}} \rightarrow \hat{\boldsymbol{p}}(\boldsymbol{r}) - \dfrac{e}{c}\boldsymbol{A}(\boldsymbol{r},t)$, we obtain the full Hamiltonian $\hat{\mathcal{H}}$:

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{\mathcal{H}}'(t) = \frac{1}{2m}(\hat{\boldsymbol{p}}(\boldsymbol{r}) - \frac{e}{c}\boldsymbol{A}(\boldsymbol{r},t))^2 + \hat{V}(\boldsymbol{r}), \tag{1.17}$$

where $\boldsymbol{A}(\boldsymbol{r},t) = \boldsymbol{A}(\boldsymbol{r})e^{\mathbb{i}(\boldsymbol{Kr}-\omega t)}$ is the vector potential of the electromagnetic radiation, e is the elementary charge, c is the speed of light in vacuum, $\boldsymbol{K}$ is the wave vector of the light field and ω is its angular frequency. In the weak field approximation, $\boldsymbol{A}^2(\boldsymbol{r},t) \approx 0$, the Hamiltonian can be rewritten as

$$\hat{\mathcal{H}} = \frac{\hat{\boldsymbol{p}}^2(\boldsymbol{r})}{2m} - \frac{e}{mc}\boldsymbol{A}(\boldsymbol{r},t)\cdot\hat{\boldsymbol{p}}(\boldsymbol{r}) + \cancelto{0}{\frac{e^2}{2mc^2}\boldsymbol{A}^2(\boldsymbol{r},t)} + \hat{V}(\boldsymbol{r}) = \hat{\mathcal{H}}_0 + \hat{\mathcal{H}}'(t), \tag{1.18}$$

where we use the Coulomb gauge div$\boldsymbol{A} = 0$, which allows $\hat{\boldsymbol{p}}$ to commute[4] with $\boldsymbol{A}$. The second term $\hat{\mathcal{H}}' = -\frac{e}{mc}\boldsymbol{A}(\boldsymbol{r},t)\cdot\hat{\boldsymbol{p}}(\boldsymbol{r})$ corresponds thereby to the perturbation Hamiltonian. Its non-oscillatory part then equates to $\hat{H}' = -\frac{e}{mc}\boldsymbol{A}(\boldsymbol{r})\cdot\hat{\boldsymbol{p}}(\boldsymbol{r})$. If we assume that the light wavelength λ is large compared to atomic dimensions a_0, i.e. $\boldsymbol{Kr} \sim 2\pi\dfrac{a_0}{\lambda} \ll 1$, then the vector potential looses its spatial dependence $\boldsymbol{A}(\boldsymbol{r}) = \boldsymbol{A} \approx$ const and the exponential factor can be expanded in Maclaurin series $e^{i\boldsymbol{Kr}} = 1 + \boldsymbol{Kr} + ... \approx 1$, where we keep only the first term. This is the so-called dipole approximation, in which the transition probability equates to

$$w_{\mathrm{fi}} = \frac{2e^2\pi}{\hbar m^2c^2}\,|\langle\psi_{\mathrm{f}}\,|\,\boldsymbol{A}\cdot\hat{\boldsymbol{p}}(\boldsymbol{r})\,|\,\psi_{\mathrm{i}}\rangle|^2\,\delta(E_{\mathrm{f}} - E_{\mathrm{i}}^{(0)} - \hbar\omega). \tag{1.19}$$

When the photoelectrons escape into vacuum, they have to overcome the work function of the sample and endure refraction at the surface, upon which only the parallel component of their wave vector is conserved. The expression for the photoelectron current then reads as

$$\begin{aligned} I(\theta,\varphi;E_{\mathrm{kin}}) \quad &\propto \quad \sum_i |\langle\psi_{\mathrm{f}}(\theta,\varphi;E_{\mathrm{kin}})|\,\boldsymbol{A}\cdot\boldsymbol{p}\,|\psi_i\rangle|^2 \\ &\qquad \times\delta(E_{\mathrm{kin}} + \Phi_0 - E_i - \hbar\omega)\times\delta(\boldsymbol{k}_{\mathrm{f}\|} - \boldsymbol{k}_{i\|} - \boldsymbol{G}_{\|}), \end{aligned} \tag{1.20}$$

where the photocurrent is given by a sum over all transitions from the occupied initial states $|\psi_i\rangle$ into the final state $|\psi_{\mathrm{f}}\rangle$ characterised by the direction (θ,φ) and kinetic energy E_{kin} of emitted photoelectrons. The polar and azimuthal emission angles are denoted by θ and φ, respectively. The delta functions comprising photon energy $\hbar\omega$, sample work function Φ_0, momentum components of the initial and final states, $\boldsymbol{k}_{\mathrm{i}\|}$ and $\boldsymbol{k}_{\mathrm{f}\|}$, parallel to the surface, respectively, and reciprocal lattice vector $\boldsymbol{G}_{\|}$ ensure energy and momentum conservation during photoemission. Because energy and the photoelectron momentum component parallel to the surface are conserved, angle-resolved photoemission can be used to probe the band structure of solids.

1.3 PHOTOEMISSION FROM MOLECULAR ADSORBATES

1.3.1 *Plane wave approximation*

Within the framework of the one-step model, the photoemission from molecular adsorbates can be described using Eq. 1.20 as a single coherent process of electron excitation from an initial state $|\psi_i\rangle$ into the final state $|\psi_{\mathrm{f}}\rangle$. The most difficult part in the evaluation of the expression for the photocurrent is a proper description of the photoelectron final state $|\psi_{\mathrm{f}}\rangle$. In the simplest case, the final state interaction between the photoemitted electron and the photohole can be neglected and the final state electron wave function can be described by a plane wave: $|\psi_{\mathrm{f}}\rangle \propto e^{\mathbb{i}\boldsymbol{k}_{\mathrm{f}}\boldsymbol{r}}$. This approach was suggested by Gadzuk [30] and was successfully used to explain photoemission distributions from atoms and small molecules adsorbed on surfaces.

Recently, Puschnig et al. [1] suggested that the approximation of the final state by a plane wave is particularly beneficial for the description of photoemission from molecular adsorbates. In this case, the

---

[4] $\hat{\boldsymbol{p}}\cdot\boldsymbol{A} - \boldsymbol{A}\cdot\hat{\boldsymbol{p}} = \frac{\hbar}{\mathbb{i}}\nabla\cdot\boldsymbol{A}$ and if $\nabla\cdot\boldsymbol{A} = 0$, operators $\hat{\boldsymbol{p}}$ and $\boldsymbol{A}$ commute: $\hat{\boldsymbol{p}}\cdot\boldsymbol{A} = \boldsymbol{A}\cdot\hat{\boldsymbol{p}}$

photoemission intensity distribution, acquired from one particular molecular initial state i, becomes proportional to the squared modulus of the Fourier transform of the initial state wave function, modulated by the $|\boldsymbol{A}\cdot\boldsymbol{k}_\mathrm{f}|^2$ factor:

$$\begin{aligned} I \propto \left|\langle e^{\mathbb{i}\boldsymbol{k}_\mathrm{f}\boldsymbol{r}}|\,\boldsymbol{A}\cdot\frac{\hbar}{\mathbb{i}}\nabla\,|\psi_\mathrm{i}(\boldsymbol{r})\rangle\right|^2 &= \left|\int e^{-\mathbb{i}\boldsymbol{k}_\mathrm{f}\boldsymbol{r}}\boldsymbol{A}\cdot\frac{\hbar}{\mathbb{i}}\nabla\psi_\mathrm{i}(\boldsymbol{r})d\boldsymbol{r}\right|^2 = \\ = \left|\int \psi_\mathrm{i}^*(\boldsymbol{r})\boldsymbol{A}\cdot\frac{\hbar}{\mathbb{i}}\nabla e^{\mathbb{i}\boldsymbol{k}_\mathrm{f}\boldsymbol{r}}d\boldsymbol{r}\right|^2 &= \hbar^2|\boldsymbol{A}\cdot\boldsymbol{k}_\mathrm{f}|^2\left|\int \psi_\mathrm{i}^*(\boldsymbol{r})e^{\mathbb{i}\boldsymbol{k}_\mathrm{f}\boldsymbol{r}}d\boldsymbol{r}\right|^2 = \\ = \hbar^2|\boldsymbol{A}\cdot\boldsymbol{k}_\mathrm{f}|^2\left|\int \psi_\mathrm{i}(\boldsymbol{r})e^{-\mathbb{i}\boldsymbol{k}_\mathrm{f}\boldsymbol{r}}d\boldsymbol{r}\right|^2 &= \hbar^2|\boldsymbol{A}\cdot\boldsymbol{k}_\mathrm{f}|^2|\psi_\mathrm{i}(\boldsymbol{k}_\mathrm{f})|^2 \end{aligned} \tag{1.21}$$

This result implies that if the phase distribution $\arg\{\psi_\mathrm{i}(\boldsymbol{k}_\mathrm{f})\}$ in the detector plane were known as well, the inverse Fourier transform of the complex-valued distribution $\sqrt{I}e^{\mathbb{i}\arg\{\psi_\mathrm{i}(\boldsymbol{k}_\mathrm{f})\}}/|\boldsymbol{A}\cdot\boldsymbol{k}_\mathrm{f}|^2$ would yield the complex-valued distribution of the initial state wave function. However, the phase distribution in the detector plane is lost, since only the squared modulus of the amplitude, i.e. $|\psi_\mathrm{i}(\boldsymbol{k}_\mathrm{f})|^2$, can be measured experimentally. This issue constitutes the so-called phase problem: an object distribution is to be reconstructed from the modulus of its Fourier transform only. In some cases, the phase distribution can be inferred from the parity of the wave function [1] or from dichroism measurements [7]. However, this is not applicable to the most general type of problems when the phase distribution cannot be deduced from symmetry considerations. This issue was addressed by Lueftner et al [3] who employed an iterative phase retrieval procedure similar to the Gerchberg-Saxton algorithm [31]. In this procedure, one iterates back and forth between real and reciprocal spaces by computing Fourier transforms and imposing constraints in both domains. In real space, the wave function distribution is confined to a rectangular box whose size is on the order of the van der Waals size of the molecule. The absolute value of the wave function distribution outside of the confinement box is reduced by 10% at each iteration step. In reciprocal space, the computed amplitude distribution is replaced by the measured one and the phase distribution is kept. In this work, we suggest that the phase problem in orbital tomography can be solved in a more robust manner by making use of the analogy to the phase problem in coherent diffraction imaging (CDI) [8].

1.3.2 *Remarks about the validity of the plane wave approximation*

Before we discuss the means of how the phase distribution in the detector plane can be retrieved, let us make some remarks about the validity of the plane wave (PW) approximation. It was observed that the PW approximation (i) fails to describe the photoemission intensity of some large polyatomic molecules as it does not account for spherical wave effects of the outgoing photoelectron waves and (ii) that it falsely predicts zero photoemission intensity in certain experimental geometries [32, 33]. To account for these inconsistencies, the independent atomic centre (IAC) approximation was developed by Grobman [34]. In the IAC formalism, the photoemission is calculated as an independent coherent sum from individual atomic centres that build the molecule. This allows for a more precise description of the photoemission final state. Under certain conditions, the IAC approximation can be nevertheless reduced to the PW approximation. Namely, it was shown that a plane wave adequately describes the photoelectron final state if (i) the photoemission happens from molecular orbitals of the same character, e.g. p_z-orbitals of large planar molecules, (ii) molecules consist of mainly light atoms (H,C,N,O) and final state scattering effects can thus be neglected [1, 2].

2

COHERENT DIFFRACTION IMAGING

A beautiful analogy to the phase problem in photoemission can be found by looking at diffraction of light waves at matter. As the most simple example, let us consider the diffraction of a plane wave on an aperture in an opaque screen.

KIRCHHOFF INTEGRAL [35] An analytic solution to a homogeneous wave equation, describing the diffraction of the light field on an aperture in an opaque screen, is given by the so-called Kirchhoff diffraction integral. Mathematically, this integral describes the Huygens principle, which explains the wave phenomena by means of a decomposition of light waves into spherical harmonics instead of decomposing them into plane waves as it is done in Fourier analysis.

Assuming that the distance from the source to a detection point is large compared to the transverse dimensions of the aperture, the Kirchhoff integral can be written as

$$u(\boldsymbol{r}) = \frac{\mathbb{i}}{\lambda} \int_{\mathcal{A}} u(\boldsymbol{r}') \frac{e^{-\mathbb{i} K|\boldsymbol{r}\text{-}\boldsymbol{r}'|}}{|\boldsymbol{r}\text{-}\boldsymbol{r}'|} d\mathcal{A}, \tag{2.1}$$

where λ is the wavelength of light, the first term in the integral is the complex-valued incident light field[1] evaluated at the point $\boldsymbol{r}'$ and the second term represents the excited spherical wave propagated to the observed point $\boldsymbol{r}$. The integral is taken over the whole area of the aperture $\mathcal{A}$. The wave vector of the light wave in vacuum, $\boldsymbol{K}$, obeys the dispersion relation

$$\boldsymbol{K}^2(\omega) = K_x^2 + K_y^2 + K_z^2 = \frac{\omega^2}{c^2}, \tag{2.2}$$

where K_x, K_y and K_z are the wave vector components projected onto x, y and z axes of the Cartesian coordinate system, respectively. The absolute value of the wave vector is connected to the wavelength of light as $K = \frac{2\pi}{\lambda} = 2\pi q$, where the quantity q is called the spatial frequency. Typically, the Kirchhoff integral is not evaluated directly, but instead is often simplified by imposing different approximations [37].

FRESNEL AND FRAUNHOFER APPROXIMATIONS Let us consider a plane wave travelling in z direction and incident normally on an opaque screen with an aperture located in the x-y plane at $z = 0$. The field distribution right after the aperture is computed as $u(x, y, z = 0) = u_0 \cdot t(x, y)$, where u_0 is the incident light field distribution right before the aperture and $t(x, y)$ is a transmission function of the aperture, which can be complex-valued. In the following, $u(x, y, z = 0)$ is referred to as the field distribution in the object plane. In general, it is a complex valued function: $u(x, y, z = 0) = |u(x, y, z = 0)| \cdot e^{\mathbb{i}\varphi(x,y,z=0)}$, where $|u(x, y, z = 0)|$ is the amplitude and $\varphi(x, y, z = 0)$ is the phase distribution in the object plane. We are interested in the field distribution far away from the screen. Thereby, we need to propagate the field distribution $u(x, y, z = 0)$ into the half-space $z > 0$.

The propagation of optical fields in a homogeneous, isotropic, linear and source-free medium can be conveniently done by means of an angular spectrum method, which was first introduced by J. A. Ratcliffe [38] and is in detail described in the book of J. W. Goodman [39]. The Fourier transform of the field distribution in the object plane yields the initial angular spectrum, which is also often referred to as the initial spatial frequency spectrum:

$$U(K_x, K_y, z = 0) = \frac{1}{(2\pi)^2} \iint\limits_{-\infty}^{+\infty} u(x, y, z = 0) e^{-\mathbb{i}(K_x x + K_y y)} dx dy. \tag{2.3}$$

The angular spectrum in the detector plane at $z = z_s$ is obtained by a multiplication of the initial angular spectrum by a complex-valued transfer function $H(K_x, K_y; z_s) = e^{\mathbb{i} K_z z_s}$:

$$U(K_x, K_y, z_s) = U(K_x, K_y, z = 0) e^{\mathbb{i} K_z z_s}, \tag{2.4}$$

---

[1] Note that in the following we use a scalar representation of electromagnetic fields, as it is justified in the absence of currents and charges [36].

where $K_z = \sqrt{\boldsymbol{K}^2 - K_x^2 - K_y^2}$ is the z-component of the wave vector. The field distribution in the detector plane is obtained via inverse Fourier transform of the propagated spectrum:

$$u(\tilde{x},\tilde{y},z_s) = \iint\limits_{-\infty}^{+\infty} dK_x dK_y \cdot U(K_x,K_y,z=0) \cdot e^{\mathbb{i}\,K_z z_s} \cdot e^{\mathbb{i}\,(K_x\tilde{x}+K_y\tilde{y})}, \tag{2.5}$$

where $\tilde{x}$ and $\tilde{y}$ denote the Cartesian coordinates in the detector plane.

In paraxial approximation, which is also often called Fresnel approximation, we assume the angular spectrum to be narrow band, i.e. $K_x^2 + K_y^2 \ll K^2$. This corresponds to the requirement that all structural details of the field distribution in the object plane must be larger than the wavelength of the light field: $|x|, |y| \gg \lambda$. This allows us to expand K_z in Taylor series around $K_x^2 + K_y^2 = 0$ as $K_z \approx K - \frac{K_x^2+K_y^2}{2K}$. In this case, the transfer function becomes

$$H(K_x,K_y;z_s) \approx H_{\mathrm{F}}(K_x,K_y;z_s) = e^{\mathbb{i}\,K z_s} e^{-i\frac{K_x^2+K_y^2}{2K} z_s} \tag{2.6}$$

and the light field distribution in the detector plane in paraxial approximation is computed as

$$u_{\mathrm{F}}(\tilde{x},\tilde{y},z_s) = \iint\limits_{-\infty}^{+\infty} U(K_x,K_y,z=0) H_{\mathrm{F}}(K_x,K_y;z_s) e^{\mathbb{i}\,(K_x\tilde{x}+K_y\tilde{y})} dK_x dK_y = \tag{2.7}$$

$$= \iint\limits_{-\infty}^{+\infty} u(x,y,z=0) h_{\mathrm{F}}(\tilde{x}-x,\tilde{y}-y,z_s) dxdy = -\frac{\mathbb{i}\,K}{2\pi z_s} e^{\mathbb{i}\,K z_s} \iint\limits_{-\infty}^{+\infty} u(x,y,z=0) e^{\mathbb{i}\,\frac{K}{2z_s}((\tilde{x}-x)^2+(\tilde{x}-y)^2)} dxdy,$$

where the spatial response function $h_{\mathrm{F}}(\tilde{x}-x,\tilde{y}-y,z_s)$ corresponds to the Fourier transform of the transfer function $H(K_x,K_y;z_s)$ and the second equality follows from the convolution theorem.

Eq. 2.8 corresponds to the Kirchhoff diffraction integral in paraxial approximation. It can be further simplified in the limit of very large propagation distances $z_s \gg \lambda$. For that we introduce the so-called Fresnel number, $N_F = \frac{a}{\lambda}\frac{a}{z_s}$, where a is the largest size of the aperture, and impose the requirement $N_F \leq 0.1$ [37]. In this case, the exponential factor $e^{\mathbb{i}\,\frac{K}{2z_s}(x^2+y^2)}$ becomes ≈ 1. This approximation, called Fraunhofer approximation, is hence valid only in the so-called "far field", that is far away from the aperture. The field distribution in the detector plane thereby becomes

$$u_{\mathrm{FR}}(\tilde{x},\tilde{y},z_s) = -\frac{\mathbb{i}\,K}{2\pi z_s} e^{\mathbb{i}\,K z_s} e^{\mathbb{i}\,\frac{K}{2z_s}(\tilde{x}^2+\tilde{y}^2)} \iint\limits_{-\infty}^{+\infty} u(x,y,z=0) e^{-\mathbb{i}\,\left(\frac{K\tilde{x}}{z_s}x+\frac{K\tilde{y}}{z_s}y\right)} dxdy =$$

$$= -\frac{\mathbb{i}\,2\pi K}{z_s} e^{\mathbb{i}\,K z_s} e^{\mathbb{i}\,\frac{K}{2z_s}(\tilde{x}^2+\tilde{y}^2)}\; U\left(\frac{K\tilde{x}}{z_s},\frac{K\tilde{y}}{z_s},z=0\right),$$

where $U\left(\frac{K\tilde{x}}{z_s},\frac{K\tilde{y}}{z_s},z=0\right)$ is the Fourier transform of the field distribution in the object plane. Hence, in the far field approximation, in each point $\tilde{x}$ and $\tilde{y}$ in the detector plane, only one angular frequency $q_x = \frac{\tilde{x}}{\lambda z_s}$, $q_y = \frac{\tilde{y}}{\lambda z_s}$ with the spectral amplitude $U\left(2\pi\frac{\tilde{x}}{\lambda z_s}, 2\pi\frac{\tilde{y}}{\lambda z_s}, z=0\right)$ contributes to the field distribution. The intensity distribution in the far field is then given by the squared modulus of the Fourier transform of the field distribution in the object plane:

$$I_{\mathrm{FR}}(\tilde{x},\tilde{y},z_s) \sim \left|U\left(2\pi\frac{\tilde{x}}{\lambda z_s}, 2\pi\frac{y}{\lambda z_s}, z=0\right)\right|^2. \tag{2.8}$$

The fact that the field distributions in the detector and the object planes are connected via the Fourier transform forms the basis for a lensless imaging technique called coherent diffraction imaging (CDI) [8]. In conventional imaging, the image is formed by an optical lens system, which performs the Fourier transform. In contrast, in CDI, only the amplitude of the Fourier transform, $|U|$, is recorded and the unknown phase distribution in the detector plane, $\arg\{U\}$, has to be retrieved numerically, in order to be able to compute the object distribution from the inverse Fourier transform of $|U|e^{\mathbb{i}\arg\{U\}}$. The loss of the phase distribution constitutes the phase problem, the one we referred to in the context of the photoemission from molecular adsorbates in chapter 1. Mathematically, Eq. 1.21 and Eq. 2.8 are identical, differing only by some constant pre-factors. The goal in both cases is the computation of the object distribution from the amplitude of its Fourier transform only. Therefore, it is intuitively clear that the mathematical methods for the solution of the phase problem in CDI must be applicable to the solution of the phase problem in orbital tomography.

PHASE RETRIEVAL ALGORITHMS

3

> *Clearly, a very intelligent algorithm is needed to search through the possible objects to find one consistent with the Fourier modulus data.*
>
> J.C. Dainty and J.R. Fienup, 1987 [40]

3.1 THE PHASE PROBLEM

Following the discovery of x-rays by Wilhelm Conrad Röntgen in 1895 [41], Max von Laue was determined to prove their electromagnetic nature by performing an experiment, in which the x-rays could be diffracted by a periodic arrangement of slits or holes. Laue realised that a periodic crystal structure may serve as a diffraction grating for x-rays[1]. The idea of Laue was proved experimentally by his colleagues W. Friedrich and P. Knipping, who observed a diffraction pattern from a copper sulphate crystal [44]. Laue developed a supporting theory and was awarded the Noble Prize in Physics in 1914 "for his discovery of the diffraction of x-rays by crystals". Some years later, William Henry Bragg and his son Willian Lawrence Bragg suggested that if the phases of the diffraction spots were known, the inverse Fourier transform of the diffraction pattern would directly yield the atomic arrangement within the crystal. In 1952, following a conversation with Alan Turing, David Sayre published an article [45], in which he draw an analogy between the Shannon sampling theorem in communication theory [46] and the sampling requirements in the field of crystallography. Sayre stated that the phase problem for a centrosymmetric crystal would be solvable, provided the diffraction pattern could be sampled at twice the Bragg frequency [45]. However, no practical algorithms for solution of the phase problem were suggested.

The first algorithm emerged only 20 years later, when Gerchberg and Saxton published their seminal paper on the determination of the phase distribution of a function whose intensity distributions in the object and diffraction planes were known [31]. This algorithm in its original form cannot be applied to a broad class of problems, as an estimate of the object's distribution is not always available. For instance, in the field of crystallography or orbital tomography, one has access only to the intensity distribution in reciprocal space. Luckily, in 1978, James Fienup suggested a modified version of the algorithm of Gerchberg and Saxton, known as the error reduction (ER) algorithm [11]. Fienup's algorithm allowed reconstructing an object distribution solely from the modulus of its Fourier transform. Instead of using the low resolution image of the object distribution, Fienup introduced positivity and support constraints in the object domain. The positivity constraint required the object distribution to be positive. The support constraint required the object distribution to be confined within a certain region which is called support. Fienup defined the support, S, of a real-valued, non-negative function $f(x) \geq 0$, where $x \in \mathcal{E}^N$, as the smallest closed set such that the integral of $f(x)$ over the complement of S in $\mathcal{E}^N$ is zero [47]. Here, $\mathcal{E}^N$ denotes N-dimensional Euclidian space. Loosely speaking, the support is a set of points over which the object distribution is nonzero [48]. Fienup gave an upper limit for the diameter of the object distribution as one half of the diameter of the autocorrelation function [11, 47]. Although, as he admitted later, in practice, this requirement is difficult to fulfil exactly, as "in two dimensions the exact support of the object cannot in general be determined uniquely from the support of its autocorrelation and so the diameter constraint cannot be applied very tightly" [49]. Since the ER algorithm suffered from slow convergence and tendency to stagnate in local minima, some years later, Fienup developed the hybrid input-output (HIO) algorithm [49]. In 2010, Ross John Harder extended it towards complex-valued object distributions by introducing the phase-constrained version (phase-constrained hybrid input-output, PC-HIO, algorithm) [13].

---

[1] Auguste Bravais predicted the interatomic separation in crystal lattices to be on the order of an Ångström [42] and Arnold Sommerfeld estimated that the x-rays should possess extremely short wavelengths on the same order of magnitude, i.e. one third of an Ångström [43].

3.2 RECONSTRUCTION PROCEDURE

3.2.1 *Hybrid input-output and error reduction algorithms*

Both the ER and the HIO algorithms employ a similar scheme which is shown in Fig. 3.1. The only difference is the application of the object domain constraints.

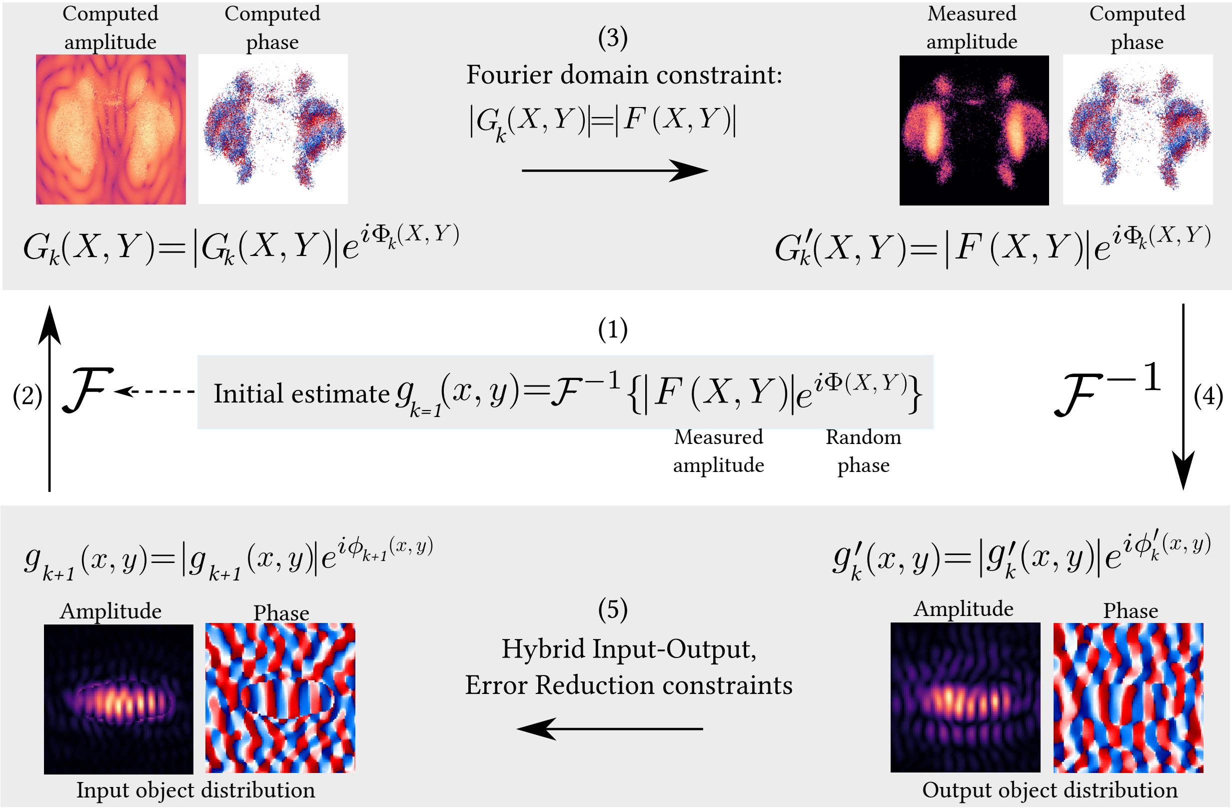


Figure 3.1: Iterative phase retrieval scheme. Adapted from [PK3].

1. In the first iteration $k = 1$, the experimental amplitude $|F(X,Y)| = \sqrt{I(X,Y)}$ is combined with a random phase distribution and the inverse Fourier transform supplies an initial input object distribution $g_{k=1}(x,y)$. Here and further in the text, (X,Y) and (x,y) denote the generalised coordinates in the detector and object planes, respectively. The random phase distribution provides an unbiased initial guess. Though, as it was stated by Gerchberg and Saxton [31], this is "... not necessary in every case and indeed there is every reason to suppose that an educated guess at the correct phase distribution would lessen the computation time required for the process to achieve an acceptable squared error". Importantly, in case of centrosymmetric distributions in the image and detector planes, one should avoid using constant phase distribution as an initial guess, because, according to Gerchberg and Saxton [31], this "will cause the algorithm to fail". The reason for this is that the "phase distribution will not change under Fourier Transformation in this case" [31].
2. By computing the Fourier transform of $g_k(x,y)$, we obtain the complex-valued distribution $G_k(X,Y) = \mathcal{F}\{g_k(x,y)\}$.
3. At this point, we satisfy the reciprocal space constraints and replace the calculated amplitude $|G_k(X,Y)|$ with the experimental amplitude $|F(X,Y)|$, while keeping the calculated phase distribution. Upon replacement, we obtain an updated complex-valued field distribution in the detector plane $G'_k(X,Y)$.
4. Inverse Fourier transform of $G'_k(X,Y)$ provides the output object distribution $g'_k(x,y)$.

5. Now, we have to satisfy the object domain constraints. In the PC-HIO algorithm [11, 13], the input object distribution for the next iteration $g_{k+1}(x,y)$ is obtained as

$$g_{k+1}(x,y) = \begin{cases} g'_k(x,y), & \text{if } (x,y) \in \gamma, \\ g_k(x,y) - \beta g'_k(x,y), & \text{if } (x,y) \notin \gamma, \end{cases} \tag{3.1}$$

where $\beta = 0.9$ is a feedback parameter and γ corresponds to a set of points which comply with the object domain constraints (belong to the support region S and have their phases within an expected range). In the ER algorithm [11], the object distribution $g_{k+1}(x,y)$ is calculated as

$$g_{k+1}(x,y) = \begin{cases} g'_k(x,y), & \text{if } (x,y) \in \gamma, \\ 0, & \text{if } (x,y) \notin \gamma, \end{cases} \tag{3.2}$$

where γ fulfils the same criteria as in the PC-HIO algorithm.

Typically, the reconstruction procedure consists of a combination of the PC-HIO and the ER algorithms. They are used in an alternating scheme, as it helps to eliminate stagnation problems and to provide faster convergence [11, 13, 50]. As it was suggested by Fienup, a reconstruction procedure starts with several iterations (e.g. 20) of the ER algorithm [48]. These are then followed by a number of cycles of iterations, where one cycle consists of several (e.g. 10-30) iterations of the HIO algorithm accompanied by a fewer number (e.g. 5-10) of iterations of the ER algorithm [48, 49].

3.2.2 *Support constraint and shrinkwrap algorithm*

The convergence speed and the accuracy of the reconstruction depend very much on the quality of the support constraint. In case of complex-valued object distributions, this choice is particularly important, as the positivity constraint is no longer applicable. In general, reconstruction of complex-valued object distributions is more difficult than reconstruction of real-valued object distributions. For instance, the complex-valued images reconstructed by Bates and Tan with Fienup's algorithm and a loose support distribution were "barely recognizable and virtually unrecognizable" [51]. Fienup showed that the correct reconstruction in case of complex-valued object distributions is possible only when the support constraint is tight enough [48, 52] and "is one of a number of special types of support constraints". These special types include (i) supports having separated parts and (ii) supports for which the object can be reconstructed by using the recursive algorithm with latent reference points; the latter class of objects includes objects with supports whose convex hulls have no parallel sides. Simple symmetric supports such as circles or rectangles do not work well [48]. The algorithm also works much better for objects having sharp edges than for objects having tapered edges [48].

The autocorrelation of the object distribution provides a very good starting estimate of the object support, as it was initially suggested by Fienup [47]. However, it might be very difficult, if not impossible, to obtain a tight object support distribution solely from the autocorrelation as only upper bounds on the support can be determined [11, 47]. In 2003, Marchesini proposed the shrinkwrap algorithm [12], in which the tight support distribution is found gradually in the course of the reconstruction procedure. In the shrinkwrap algorithm, the initial estimate of the object support is obtained from the autocorrelation of the object distribution by (i) computing the inverse Fourier transform of the experimental diffraction pattern $I(X,Y)$, (ii) convolving it with a Gaussian function and (iii) applying a threshold so that the pixel values below and above the threshold are set to 1 and 0, respectively. The width of the Gaussian function and the threshold level may vary from case to case, as they depend on many parameters, such as noise level, the degree of the object confinement, etc. The threshold value and the Gaussian width are chosen empirically so that no part of the reconstructed pattern is truncated. The thresholding procedure provides us with a binary mask which is used as the support constraint in object space.

In the end of each PC-HIO–ER cycle, i.e. in the last iteration of the ER algorithm, the output object distribution $g'_k(x,y)$ is used to update the object support. The updated support distribution is obtained by convolving it with a Gaussian function and setting a threshold. The width of the Gaussian is reduced at every support update by approximately 1% compared to its preceding value as it is conventionally done in the shrinkwrap algorithm [12].

3.2.3 *Error metrics*

The phase retrieval algorithm continues until a Fourier transform pair that satisfies the constraints in both domains is found [49]. This pair is regarded as a solution and the algorithm is stopped. As a measure of convergence of the algorithm towards the solution, Fienup defined a normalised root-mean-square (RMS) error metric in reciprocal space E_R. This metric computes the mismatch between the iterated and the experimental amplitudes [11]:

$$E_\mathrm{R} = \sqrt{\frac{\sum_{\text{all pixels}} ||F(X,Y)| - |G_\mathrm{it}(X,Y)||^2}{\sum_{\text{all pixels}} |F(X,Y)|^2}}, \tag{3.3}$$

where $|F(X,Y)|$ is the experimental amplitude, $|G_\mathrm{it}(X,Y)|$ is the iteratively obtained amplitude.

The equivalency between the ER algorithm and a steepest-descent gradient search algorithm, applied for the general problem of arbitrary data and constraints, "gives a proof of convergence in the weak sense that the RMS error could only decrease or stay the same at each iteration" [49]. It should be noted that the RMS error provided by Eq. 3.3 is meaningless for the PC-HIO algorithm, as the input object distribution $g(x,y)$ does not longer represent an estimate of the object distribution [49, 53]. Instead, it should be thought of as "the driving function for the next output" $g'(x,y)$ [49]. Thereby a more meaningful error for the PC-HIO algorithm would be an object domain error E_0 [49, 53] defined as

$$E_0 = \sqrt{\frac{\sum_{\notin\gamma} |g'_k(x,y)|^2}{\sum_{\text{all pixels}} |g'_k(x,y)|^2}}, \tag{3.4}$$

where γ denotes the object domain constraints and the summation is done over all pixels that these constraints violate. The relation between the value of E_0 and the visual quality of the reconstructed image distribution is not fully understood though. Counterintuitively, we observed that the quality of the reconstructed object distribution may get significantly improved as the error metric E_0 increases with each iteration, as it is often the case [49]. For this reason, we did not use E_0 in this work[2]. Instead, we estimate the quality of the reconstructed image distribution by recording the RMS error E_R at the end of each ER cycle, as it was originally recommended by Fienup [49, 53] and as it is typically done by the others [13, 55–58]. The final reconstruction is obtained by averaging over multiple reconstructions obtained in independent reconstruction runs with random initial phase distributions [5, 54, 58–61].

3.3 REMARKS ON FOURIER TRANSFORM

3.3.1 *Discrete Fourier transform*

Previously, we assumed that the signal distributions are continuous and infinitely extended. However, in reality, they are not continuous, but are sampled at a finite number of pixels as governed by the discrete nature of detectors. This means that the continuous Fourier transform must be replaced with the discrete Fourier transform. The transition from continuous Fourier transform to the discrete Fourier transform is done by sampling the signal distributions in object and reciprocal spaces in such a way that they become periodic functions of coordinate and spatial frequency, respectively, with one period being equal to N samples in each of the spaces.[3] Given a one-dimensional function g in the object domain, discretised in a way that N samples constitute one period of a periodic waveform, the discrete Fourier transform of this function is then a special case of the continuous Fourier transform and is given by N samples as

$$G\,(m\Delta q) = \sum_{n=0}^{N-1} g(n\Delta r) e^{-\mathbb{i}2\pi mn/N}, \tag{3.5}$$

where $n = 0,1,2,...,N-1$ and $m = 0,1,2,...,N-1$ are the sample numbers in real and reciprocal spaces, respectively. The sampling intervals in the object and in the reciprocal domains, Δr and Δq, respectively, are connected as $\Delta r \Delta q = \frac{1}{N}$. Note that here r represents a spatial coordinate measured in m and q is a

---

[2] Other authors [9, 52, 54] do monitor the reconstruction progress with the real space error metric

[3] For succinct derivation of the discrete Fourier transform the reader is referred to the book of E. O. Brigham. [62]

spatial frequency measured in m^{-1}. The pixel size in terms of a wave number[4], $k = 2\pi q$, measured in $\mathrm{rad} \cdot \mathrm{m}^{-1}$, can thereby be computed as $\Delta k = \frac{2\pi}{N\Delta r}$. The inverse discrete Fourier transform is given by

$$g(n\Delta r) = \frac{1}{N}\sum_{m=0}^{N-1} G\left(m\Delta q\right) e^{\mathrm{i}2\pi mn/N}, \tag{3.6}$$

where the same notations as in Eq. 3.5 are used.

For a two-dimensional function g, assuming the equal sampling rates Δr and Δq and the equal number of pixels N in both dimensions, the forward discrete Fourier transform can be written as

$$G\left(m_1\Delta q, m_2\Delta q\right) = \sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1} g(n_1\Delta r, n_2\Delta r) e^{-\mathrm{i}2\pi(m_1n_1+m_2n_2)/N}, \tag{3.7}$$

where $n_1, n_2 = 0, 1, 2, ..., N-1$ are the sample numbers in real space and $m_1, m_2 = 0, 1, 2, ..., N-1$ are the sample numbers in reciprocal space.

The inverse discrete Fourier transform is written as

$$g(n_1\Delta r, n_2\Delta r) = \frac{1}{N^2}\sum_{m_1=0}^{N-1}\sum_{m_2=0}^{N-1} G\left(m_1\Delta q, m_2\Delta q\right) e^{\mathrm{i}2\pi(m_1n_1+m_2n_2)/N}, \tag{3.8}$$

where the same notations as in Eq. 3.7 are used.

3.3.2 *Centred Fourier transform*

Discrete Fourier transforms defined by Eq. 3.5, 3.6, 3.7 and 3.8 assume that the signal coordinate systems are not centred in the computational domain. In other words, it was assumed that the values of the discrete signals are given for non-negative values of space coordinates and that the origin of the coordinate system lies, in the two-dimensional case, in the lower left corner of the computational domain given by pixel numbers $(0, 0)$. In orbital tomography and CDI, signal distributions in reciprocal space are centred at the origin of the reciprocal coordinate system. This means that for the machine computation of the Fourier transforms, Eq. 3.7 and 3.8 must be modified to ensure that the signal distributions are located in the middle of the computational domains at $(N/2, N/2)$ if these are sampled at N×N pixels. The expressions for the centred discrete Fourier transforms used in this work were derived by Latychevskaia et al. [63].

The centred discrete Fourier transform in two dimensions:

$$G\left(m_1\Delta q, m_2\Delta q\right) = e^{\mathrm{i}\pi(m_1+m_2)}\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1} g(n_1\Delta r, n_2\Delta r) e^{-\mathrm{i}2\pi(m_1n_1+m_2n_2)/N} e^{\mathrm{i}\pi(n_1+n_2)}. \tag{3.9}$$

The centred inverse discrete Fourier transform in two dimensions:

$$g(n_1\Delta r, n_2\Delta r) = \frac{1}{N^2} e^{-\mathrm{i}\pi(n_1+n_2)}\sum_{m_1=0}^{N-1}\sum_{m_2=0}^{N-1} G\left(m_1\Delta q, m_2\Delta q\right) e^{\mathrm{i}2\pi(m_1n_1+m_2n_2)/N} e^{-\mathrm{i}\pi(m_1+m_2)}. \tag{3.10}$$

In machine computation, the Fourier transforms are computed using the Fast Fourier transform algorithm (FFT). The FFT algorithm allows for a particularly rapid computation of the discrete Fourier transform. Curious readers are invited to read the book of E. O. Brigham [62] or the original paper about the FFT algorithm by Cooley and Tukey [64] to learn more about its beautiful intricacies.

3.4 OVERSAMPLING REQUIREMENTS

While solving the phase problem with the iterative algorithms, it must be kept in mind that essentially one solves a system of equations, which connect an unknown complex-valued object distribution f with

---

[4] It is a common practice, to omit the radians in the units of the wave number: $\mathrm{rad} \cdot \mathrm{m}^{-1} \to 1 \cdot \mathrm{m}^{-1}$. For the sake of being consistent with the most of the literature on orbital tomography and CDI, we adhere to the same convention here and write m^{-1} for the unit of the wave number throughout the whole following text, while we still imply that they are measured in $\mathrm{rad} \cdot \mathrm{m}^{-1}$.

the amplitude of its Fourier transform, i.e. $|F| = |\mathcal{F}\{f\}|$, where $\mathcal{F}$ denotes the Fourier transform. Given an $N \times N$ pixel sampled amplitude distribution in reciprocal space

$$|F(m_1\Delta q, m_2\Delta q)| = \left| \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} f(n_1\Delta r, n_2\Delta r) e^{-\mathbb{i}2\pi(m_1n_1+m_2n_2)/N} \right|, \tag{3.11}$$

where $n_1, n_2 = 0, 1, 2, ..., N-1$ are the sample numbers in real space and $m_1, m_2 = 0, 1, 2, ..., N-1$ are the sample numbers in reciprocal space, we obtain a set of N^2 equations, which have to be solved in order to find both the amplitude and phase distribution of f. Since f is complex-valued, there are $2 \cdot N^2$ unknown variables. Obviously, in order to obtain a non-trivial solution of the system of equations Eq. 3.11, the number of unknown variables, i.e. the number of pixels that contain the amplitude and phase values of f, must be at least equal to the number of equations, i.e. to the number of pixels that contain values of the amplitude distribution in reciprocal space $|F|$. This requirement for the solution of the phase problem is called the oversampling condition and was introduced by Miao et al. [9, 10]. In practice, it is fulfilled by a dense sampling of the experimental intensity distribution, so that the object distribution has a finite support surrounded by a zero-padded region. Miao et al. [9] defined the degree of oversampling by introducing the so-called oversampling ratio as

$$\sigma = \frac{N_{\text{total}}}{N_{\text{unknown}}}, \tag{3.12}$$

where N_{total} is the total number of pixels and N_{unknown} is the number of pixels with unknown values. For a complex-valued two-dimensional object distribution, σ must exceed 2 [9]. In each dimension of a two-dimensional data set, we can define a linear oversampling ratio

$$\varnothing = \frac{N\Delta r}{a}, \tag{3.13}$$

where N is the linear number of pixels, Δr is the size of the pixel in the object domain and a is the largest extent of the object distribution. The oversampling requirement then corresponds to $\varnothing > \sqrt{2}$ [9] in each dimension for a two-dimensional square object distribution.

3.5 REMARKS ABOUT UNIQUENESS OF THE RECONSTRUCTION

It was shown that the reconstruction of two dimensional sampled objects is almost always unique (in the absence of noise) [65]. Of course, a reconstructed image distribution may always contain inherent ambiguities such as $f(x, y)$, $e^{i\theta_c} f(x - x_0, y - y_0)$ and $e^{i\theta_c} f^*(-x - x_0, -y - y_0)$, where θ_c is a constant phase, (x_0, y_0) is an arbitrary offset and the asterisk denotes complex conjugation. This is because all of them have the same Fourier modulus. If we treat these ambiguities (phase shift, translation and conjugate image) as only ambiguities, then we agree that we reached the solution and it is unique [48].

All phase retrieval algorithms described in this chapter were implemented in MATLAB by the author of the dissertation.

4

MATERIALS, METHODS AND EQUIPMENT

> *Experiment is the final judge of theories.*
>
> *— Louis de Broglie, Nobel Lecture 1929*

4.1 ANGLE-RESOLVED PHOTOELECTRON SPECTROSCOPY

The primary experimental tool used in this work is angle-resolved photoelectron spectroscopy (ARPES) [25, 66]. In ARPES, the photoelectron distributions are recored at a broad range of polar and azimuthal angles by either rotating the sample with a suitable manipulator or by using an angle-resolving electron analyser with a large acceptance angle. By acquiring the photoelectron angular distributions at different kinetic energies, one can thereby map the electronic band structure of solids.

In this work, ARPES data were recorded using two apparatuses: (i) the mobile surface science end-station *WALküre*, located in the LaserLab at the University of Zurich (UZH), and (ii) the photoemission electron microscope (PEEM) operating in momentum mode at the NanoESCA beamline at the synchrotron radiation facility Elettra [67] in Trieste, Italy. The *WALküre* end-station was used to validate the preparation of molecular adsorbate films and for some of the time-resolved measurements. The acquisition of ARPES data for orbital tomography was performed using the PEEM apparatus. Time-resolved measurements were done at the ARTEMIS beamline [68] located at the Central Laser Facility of the Rutherford Appleton Laboratory in the United Kingdom.

WALKÜRE APPARATUS *WALküre* was built as a mobile surface science end-station for use at different light sources [69, 70]. The end-station consists of a preparation chamber (PC) located on top of an analysis chamber (AC) as it shown in Fig. 4.1(A). The samples enter the PC via an entry lock. The PC hosts a number of preparation and characterisation devices, such as (i) an ion sputter gun (SPECS, Germany) for bombardment of the sample surfaces with argon ions, in order to clean them prior to the acquisition of ARPES data or evaporation of adsorbates; (ii) two colour pyrometer (Maurer, Germany) to measure the sample temperature during annealing; (iii) Knudsen-type evaporators to prepare molecular adsorbates; (iv) a low energy electron diffraction (LEED) device (OCI Microengineering, Canada) to check the quality of molecular films or single crystal surfaces; (v) a quadrupole mass spectrometer (Pfeiffer Vacuum, Switzerland) to check the gas composition in the chamber or the quality of the Argon gas used for sputtering. The sample is mounted on a Swiss top hat type sample holder that resides on a modified VG Scienta Omniax manipulator mounted on top of the PC. The sample can be moved in x,y and z directions as well as rotated by 360° around the manipulator axis and the sample normal. The sample movement is controlled either manually or automatically with the help of stepper motors implemented in Labview software.

Acquisition of the ARPES data is done in the AC, which is equipped with two continuous wave (CW) light sources and a hemispherical electron analyser. The light sources are: (i) a twin anode XR3E2 x-ray source (Thermo VG Scientific) with two anodes (Mg K_α with the photon energy $h\nu = 1253.4$ eV and Si K_α with the photon energy $h\nu = 1739.4$ eV) and (ii) a Helium discharge lamp HIS13 (Scienta Omicron, Germany). By varying pressure and discharge current, one can achieve emission at two primary emission lines: He I_α with the photon energy $h\nu = 21.2$ eV and He II_α with the photon energy $h\nu = 40.8$ eV. Photoelectrons are detected with a hemispherical electron analyser PHOIBOS 150 WAL (SPECS, Germany). The sample is placed in front of the analyser at the working distance of 27.75 mm from the wide angle electrostatic lens that collects the photoelectrons emitted over a cone of up to ±30° in a volume of about 1mm^3 and focuses them at the entrance slit of a hemispherical unit. The potential difference between the outer and inner hemisphere sets the pass energy of the photoelectrons, i.e. the kinetic energy of the photoelectrons that follow the circular path between the hemispheres and are transmitted to the detector side. Faster electrons move at higher radii. In the end, only electrons with kinetic energies within the window of 12% of the pass energy reach two microchannel plates (MCPs). The MCPs amplify the electron signal by a factor of $10^7..10^8$ and the amplified electron distribution impinges onto the phosphorous screen. The pattern on the screen is then recorded by a CCD camera (Pixelfly, Germany) sampled at 1392×1024 pixels. In essence, a single CCD snapshot represents an angular (momentum) dispersion of the photoelectrons along the vertical axis and an energy dispersion along the horizontal axis. The schematic of the electrons trajectories in the analyser is shown in Fig. 4.1(B).

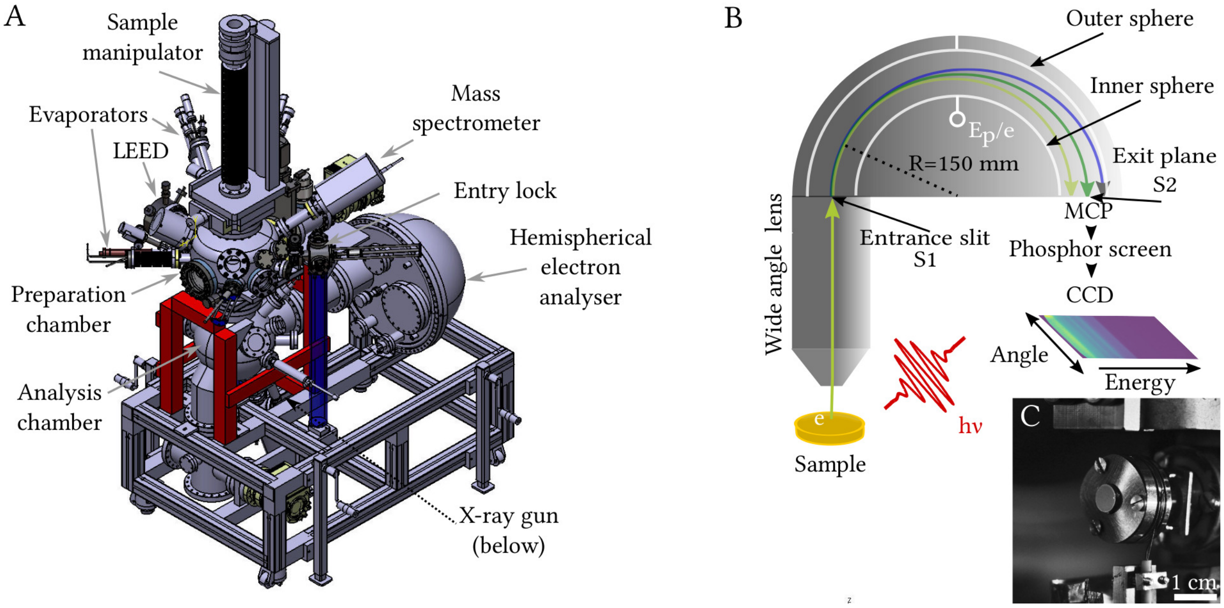


Figure 4.1: (A) – *WALküre* apparatus in the Laserlab at the University of Zurich. (B) – Electron trajectories inside the hemispherical electron analyser. Electrons ejected from the sample propagate towards the entrance slit $S1$. Electrons following the green trajectory have their energy equal to the pass energy E_p. Electrons following the light green and blue trajectories have their energies lower and higher than the pass energy, respectively and arrive at different locations at the exit plane $S2$. The electron signal distribution is amplified by MCPs and is recorded by a CCD camera. (C) – Photo of a Ag(110) single crystal residing on the sample manipulator inside the preparation chamber.

PEEM SETUP In principle, photoelectron distributions (PADs) for orbital tomography could be acquired with the *WALküre* end-station. However, apart from the apparatus capable of recording photoelectron distributions at a large range of electron emission angles, a light source with the photon energies in the order of 25–35 eV was needed as well.

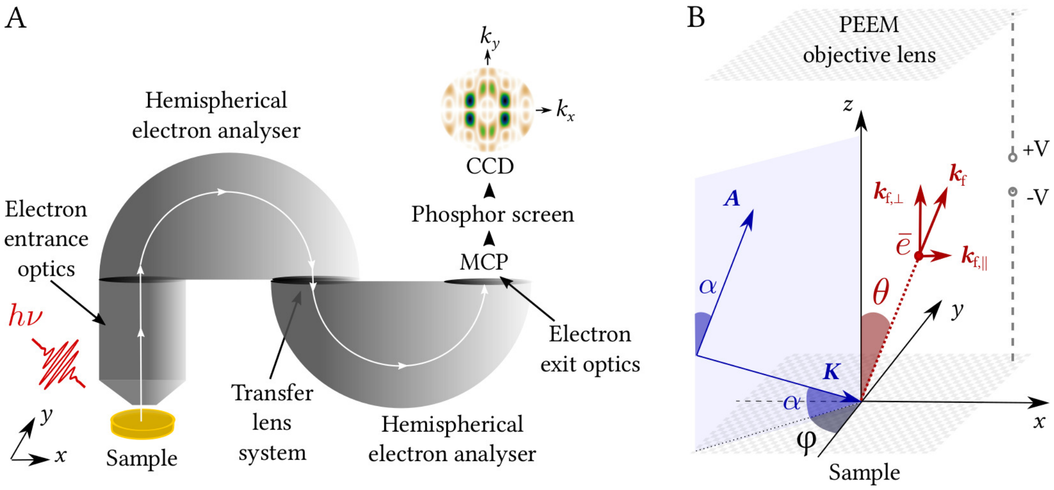


Figure 4.2: (A) – Schematic of the PEEM setup. (B) – Experimental geometry in PEEM. The 35 eV p-polarised light was incident on the sample at a grazing angle of $\alpha = 25^\circ$. The azimuthal angle between the plane of incidence and the $[1\bar{1}0]$ high symmetry direction of the Ag(110) crystal was $\varphi = 5^\circ$. The vector potential and wave vector of light are denoted by $\boldsymbol{A}$ and $\boldsymbol{K}$, respectively. The photoelectrons were collected by the PEEM objective lens. $\boldsymbol{k}_{\mathrm{f},\parallel}$ and $\boldsymbol{k}_{\mathrm{f},\perp}$ denote parallel and normal components of the final state wave vector of the photoelectrons. Adapted from [PK3] and [PK2], respectively.

This combination was provided at the NanoESCA beamline [67] of the Elettra synchrotron radiation facility (Trieste, Italy). The NanoESCA beamline is equipped with a PEEM [71, 72]. The setup of the PEEM and the experimental geometry are shown in Fig. 4.2. The 35 eV p-polarised light is incident on the sample at a grazing angle of $\alpha = 25°$. The azimuthal angle between the plane of incidence and the $[1\bar{1}0]$ high symmetry direction of the Ag(110) crystal is $\varphi = 5°$. The vector potential and the wave vector of light are denoted by $\boldsymbol{A}$ and $\boldsymbol{K}$, respectively. The photoelectrons are collected by the PEEM objective lens within the field of view of $\pm$ 2 Å^{-1}.

4.2 MOLECULES

We studied three types of organic molecules: pentacene, pyrphyrin (Pyr) and cobalt pyrphyrin (CoPyr).

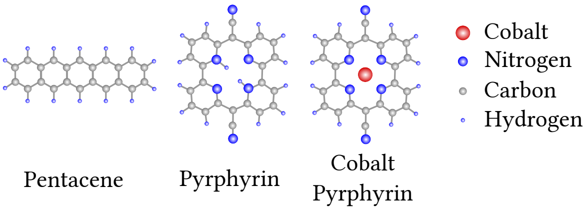


Figure 4.3: Molecules used in this work

Pentacene is a polycyclic aromatic hydrocarbon consisting of five fused benzene rings. In this work, a well-ordered sub-monolayer of pentacene molecules adsorbed on the Ag(110) substrate served as a model system for orbital tomography. The photoemission data from a pentacene sub-monolayer were acquired by Schöll and coworkers (University of Würzburg) at the NanoESCA beamline at Elettra synchrotron (Trieste, Italy) and were provided to us for validation of the phase retrieval procedure [PK2, PK3] as described in chapters 5 and 6. In the UZH lab, we prepared bilayers of pentacene on Ag(110) to study electron dynamics with time-resolved ARPES at UZH and at the ARTEMIS beamline at the Central Laser Facility (United Kingdom). Pentacene powder was acquired commercially (Sigma Aldrich, Germany).

Another type of molecules studied in this work is the Pyr molecule which is a bipyridine-based macrocycle with four pyridine rings. It was first synthesised by Ogawa et al. in 1984 [73]. CoPyr is a metalated version of Pyr: two hydrogen atoms in the middle of the molecule are replaced by one cobalt atom. In this work, ARPES data from Pyr and CoPyr on Ag(110) were acquired during a beamtime at the NanoESCA beamline at Elettra synchrotron (Trieste, Italy) with the purpose of reconstruction of molecular orbitals with phase retrieval algorithms. Pyr and CoPyr molecules used in this work were synthesised in the group of R. Alberto (UZH) [74].

Prior to the molecule deposition onto the single crystal substrate, the surface of the latter has to be cleaned. The Ag(110) used in this work was prepared in ultra high vacuum by alternating cycles of sputtering with Ar$^+$ ions and annealing at high temperatures. Sputtering was typically performed at the pressure of $3 \cdot 10^{-6}$mbar, 500 eV bias voltage, $\approx$ 540 eV discharge voltage, 4.4 mA discharge current and $\approx$ 2 μA sample current. Usually, two cycles of sputter-anneal (20 min + 10 min) were enough to ensure good surface quality. The deposition of molecules onto the Ag(110) substrate was done via a self-assembly process. Upon effusion from a home-built Knudsen cell type evaporator, molecules ordered themselves on the surface of the single crystal. Typical evaporation parameters are listed in Table 4.1.

MOLECULE	CHEMICAL STRUCTURE	EVAPORATION PARAMETERS
Pentacene	$C_{22}H_{14}$	$T_{\text{evap}} \approx 200°C$, 1ML/12 min
Pyrphyrin (Pyr)	$C_{24}N_6H_{14}$	$T_{\text{evap}} \approx 300°C$, 1ML/40 min
Cobalt Pyrphyrin (CoPyr)	$CoC_{24}N_6H_{12}$	$T_{\text{evap}} \approx 330°C$, 1ML/40 min

Table 4.1: Molecules used in this work

4.3 LOW ENERGY ELECTRON DIFFRACTION

In 1924, French physicist Louis de Broglie, made a hypothesis that electrons share the same wave-particle duality properties as quanta of light [75, 76]. He postulated that their characteristic wavelength, which was later called the de Broglie wavelength, is connected to their momentum p via the Planck's constant h as $\lambda_{\text{dB}} = h/p$. The first experimental evidence of this hypothesis dates back to 1920's, when American physicists Davisson and Germer performed a series of experiments [77–80], in which they revealed that the electrons can indeed be diffracted by a single-crystal of nickel just like visible light can be diffracted by an optical grating or x-rays can be diffracted by a single crystal[1]. They had shown that the electron wavelengths, calculated using the optical plane grating formula, satisfy the de Broglie relation and that the disagreement with the electron wavelengths calculated using the Bragg formula could be explained by assuming the refractive index of the crystal n at the given electron energies to be different from unity, unlike in case of the x-rays, for which it is ≈ 1 [80, 82]. Since the inelastic mean free path of electrons is on the order of a few atomic layers, this makes diffraction of electrons a perfect surface-sensitive tool. Thus, in the early 1960s, low energy electron diffraction (LEED) emerged as a technique for analysis of surfaces and surface confined structures under high vacuum conditions.

4.3.1 *LEED setup*

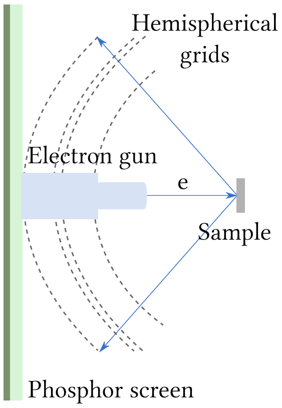


Figure 4.4: Sketch of LEED device

In this work, the LEED technique was used to determine the cleanness and the orientation of single crystal surfaces as well as to monitor the order and coverage of molecular adlayers and quantitatively characterise their surface lattices. A schematic diagram of the LEED setup is shown in Fig. 4.4. We employed LEED Model BDL 800 from OCI Microengineering, Canada. In this device, the sample is placed in front of an electron gun and the incident electron beam ($\approx$ 1–2 μA), generated by a thoriated tungsten filament, is focused by the electrostatic lenses onto the surface of a sample. Depending on the type of the sample, the electron energies are tuned between 15–50 eV for molecular adlayers and between 40–100 eV for single crystals. Upon interaction with the sample, the scattered electrons propagate back towards the flat phosphorescent screen (fused silica coated with indium-tin oxide and the green P31 phosphor ZnS:Ag:Cu). Most of the inelastically scattered electrons are high-pass filtered by the hemispherical grid. The elastically scattered electrons and the remaining inelastic ones form the LEED pattern that is recorded with the CCD camera (Model 270 XS, PCO Pixelfly, Germany) placed in front of a view port. The whole device operates in ultra-high vacuum at $\approx 10^{-9}..10^{-10}$ mbar and is magnetically shielded by a Mu-metal tube.

4.3.2 *Correction of LEED patterns*

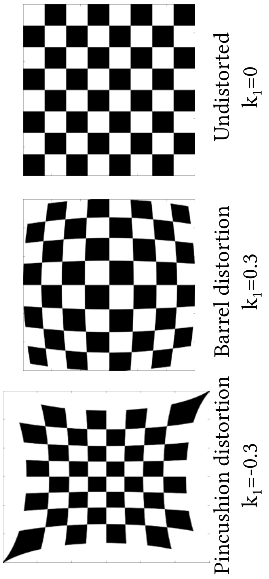


Figure 4.5: Some types of lens distortion

In order to quantitatively characterise LEED patterns recorded with the CCD camera, we had to correct (i) distortions introduced by the objective (Navitar, F0.95/17 mm) and (ii) distortions due to the projection onto the flat screen. The correction was done programmatically in MATLAB. The verification of the implemented correction algorithms was done by applying them to a LEED pattern acquired from a Si(111)7x7 surface shown in Fig. 4.6(A).

LENS CORRECTION Distortion is a monochromatic optical aberration, in which straight lines of the object space are imaged not as straight, but as curved ones [83]. Distortion occurs if the lateral magnification of the optical system becomes a function of the off-axis distance. In case of a pinhole or an ideal thin lens, the lateral magnification is constant and the distortion is absent. However, in real optical systems, the lateral magnification is constant only in paraxial approximation, i.e. near the optical axis. The real lenses are not thin, but have a finite thickness and a finite radius of curvature, which results in a radial dependency of the lateral magnification. Fig. 4.5 shows two most fundamental manifestations of distortion: barrel and pincushion distortions, in which the straight lines are bent either outwards or inwards from the image centre. Convex lenses give rise to pincushion distortions, while concave lenses lead to barrel distortions. The

---

[1] In 1937, Davisson and Thomson were jointly awarded a Nobel prize in physics for “their discovery of the interference phenomena arising when crystals are exposed to electronic beams” [81].

presence of an aperture stop also results in distortion. The aperture stop after a convex lens will result in pincushion distortion, whereas the aperture stop in front of it will lead to barrel distortion. In case of a concave lens, an opposite effect will occur [83]. The physically unobservable undistorted image coordinates, $(x_\mathrm{o}, y_\mathrm{o})$, can be related to the distorted image coordinates, $(x_\mathrm{d}, y_\mathrm{d})$ as [84]

$$\begin{aligned} x_\mathrm{o} &= x_\mathrm{d} + (x_\mathrm{d} - c_\mathrm{x})(k_1 r_\mathrm{d}^2 + k_2 r_\mathrm{d}^4), \\ y_\mathrm{o} &= y_\mathrm{d} + (y_\mathrm{d} - c_\mathrm{y})(k_1 r_\mathrm{d}^2 + k_2 r_\mathrm{d}^4), \end{aligned} \tag{4.1}$$

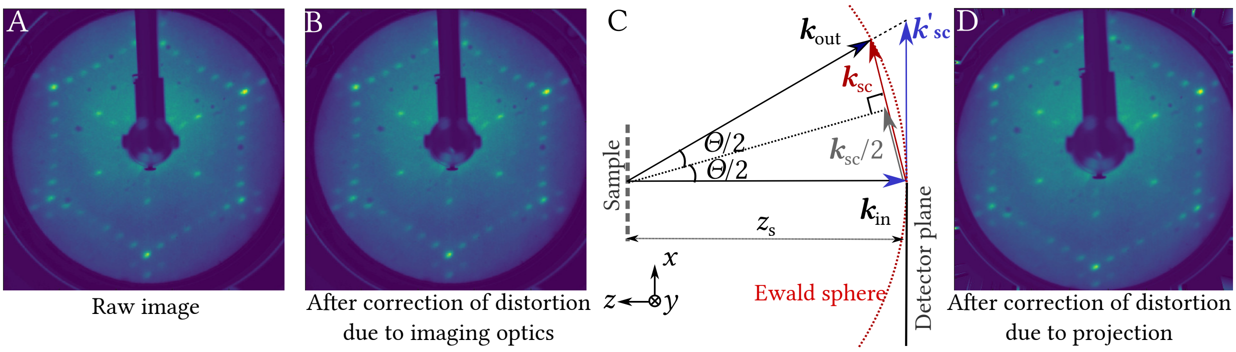
where $k_1, k_2, \ldots$ are the radial distortion coefficients, and $r_\mathrm{d}^2 = x_\mathrm{d}^2 + y_\mathrm{d}^2$ is the radial distance from an image point with coordinates $(x_\mathrm{d}, y_\mathrm{d})$ to the distortion centre located at $(c_\mathrm{x}, c_\mathrm{y})$. In most cases, it is sufficient to consider only the first two coefficients.

The distortion coefficients were obtained by calibrating the imaging system by means of the MATLAB app `cameraCalibrator`. The calibration algorithm, implemented in the `cameraCalibrator` app, was that of Zhang and its details can be found elsewhere [85]. We used a non-squared checkerboard pattern glued onto a firm flat surface of a laboratory book. The pattern was placed in focus of the camera at a distance approximately equal to that between the camera and the LEED phosphorous screen in the experiment. By rotating the pattern at angles less than 45° with respect to the imaging plane, we captured 20 images of the pattern at different orientations relative to the camera [86]. Upon entering the length of one side of a checkerboard square, $l = 25$ mm, the `cameraCalibrator` app successfully detected the checkerboard pattern in each of the added images and delivered the `cameraParameters` object containing all intrinsic and extrinsic parameters of the camera as well as the radial distortion coefficients $k_1 = -0.6076$ mm and $k_2 = 0.2437$ mm. The opposite sign of the coefficients highlights the "wavy" character of the distortion, where both barrel and pincushion types were present. The coefficients were inserted into Eq. 4.2 and the resulting LEED pattern, shown in Fig. 4.6(B), was obtained by remapping the image coordinates by means of bicubic interpolation.

DISTORTION DUE TO PROJECTION ONTO FLAT SCREEN After correction of the distortion introduced by the imaging optics, a distortion due to projection of the LEED patterns onto the flat screen was corrected. This type of distortion occurred because the wave vectors of the electrons scattered at non-zero angles Θ, $\boldsymbol{k}_\mathrm{out}$, resided on the so-called Ewald sphere, but the detected electron distribution was gnomonically [87] projected onto the flat screen as it is shown in Fig. 4.6(C). The actual scattering wave vector $\boldsymbol{k}_\mathrm{sc}$ was thereby not equal to the imaged scattering wave vector $\boldsymbol{k}'_\mathrm{sc}$:

$$|\boldsymbol{k}_\mathrm{sc}| = \frac{2|\boldsymbol{k}'_\mathrm{sc}| \sin \Theta/2}{\tan \Theta}, \tag{4.2}$$

where $\Theta = \mathrm{atan}\{\sqrt{x^2 + y^2}/z_\mathrm{s}\}$, the coordinates in the detector plane are denoted by x and y and z_s is the distance between the sample and the detector plane. Eq. 4.2 was used to remap the image coordinates of the LEED pattern in Fig. 4.6 (B) by means of bicubic interpolation and the result is shown in Fig. 4.6 (D).



Raw image
After correction of distortion due to imaging optics
After correction of distortion due to projection

Figure 4.6: (A) – LEED pattern acquired from Si(111)7x7 surface at 57 eV. (B) – The same pattern as in (A), but after the correction of the distortion due to imaging optics. (C) – Geometric relationship between the scattering wave vector, $\boldsymbol{k}_\mathrm{sc}$, and the imaged scattering wave vector, $\boldsymbol{k}'_\mathrm{sc}$. (D) – The same pattern as in (B), but after the correction of the distortion due to the projection onto the flat screen.

4.4 TIME-RESOLVED PHOTOELECTRON SPECTROSCOPY

Temporal dynamics of molecular orbitals can be studies by means of time-resolved ARPES, in which pulsed laser sources are used instead of CW laser sources for excitation. A suitable method is called two-photon photoemission (2PPE) [88–93]. In 2PPE, electrons are excited from a ground state $|0\rangle$ into an intermediate state $|1\rangle$, whose dynamics we would like to study. This is done by a laser pulse with a photon energy $h\nu_1 = E_{|1\rangle} - E_{|0\rangle}$, where $E_{|0\rangle}$ and $E_{|1\rangle}$ are the energies of the ground and the excited state, respectively. The excited state $|1\rangle$ decays on its typical relaxation time $\Delta t_{\text{relaxation}}$. The second pulse arrives at a moment in time Δt_{delay} and excites the remaining electrons from state $|1\rangle$ to a state $|2\rangle$ in vacuum. The electrons propagate towards a detector and the measured signal is proportional to the population of state $|2\rangle$ at delay time Δt_{delay}. By varying the delay time Δt_{delay} between both laser pulses, one can follow the dynamics of state $|2\rangle$. The photon energies $h\nu_1$ and $h\nu_2$ must not exceed a work function of the sample so that a single laser pulse alone cannot promote an electron into vacuum.

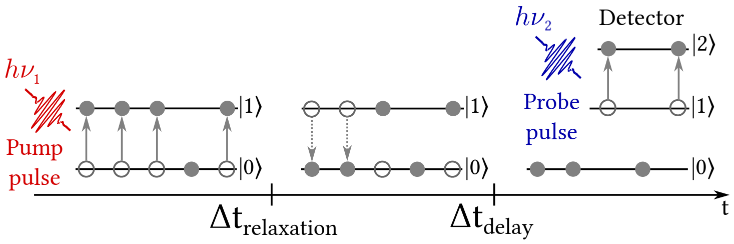


Figure 4.7: Schematic of the 2PPE experiment. Electrons are excited from an ground state $|0\rangle$ to a state $|1\rangle$ upon absorption of photons from a pump laser pulse. State $|1\rangle$ decays back into the ground state $|0\rangle$ with a decay time $\Delta t_{\text{relaxation}}$. The pump laser pulse arrives at a moment in time Δt_{delay} and excites electrons from state $|1\rangle$ to a final state $|2\rangle$. The signal measured with a detector is proportional to the population of state $|1\rangle$ at the moment Δt_{delay}.

In this work, we use the 2PPE technique to study the dynamics of unoccupied molecular states of a pentacene bilayer deposited on the Ag(110) substrate. The 2PPE experiments at the UZH lab were focused on the determination of the pump wavelength for HOMO-LUMO transition in the bilayer of pentacene on Ag(110). The actual pump-probe experiments were then performed with ARTEMIS high harmonic light source at the Central Laser Facility at the Rutherford Appleton Laboratory (United Kingdom).

4.5 LASER SETUP AT UZH

4.5.1 *Setup description*

For the 2PPE experiments at UZH, we employed a commercial laser system consisting of a Ti:Sapphire (Ti:Sa) oscillator, a Regenerative Amplifier (RegA) and an Optical Parametric Amplifier (OPA) (Coherent, USA).

DEVICE MODEL	WAVELENGTH	REP. RATE	PULSE DURATION
Ti:Sa Oscillator (MIRA, Coherent, USA)	810 nm	80 MHz	≈50 fs
Regenerative Amplifier (RegA, Coherent, USA)	≈810 nm	100 or 250 kHz	≈50 fs
OPA (Coherent, USA) - signal	480 to 700 nm	100 or 250 kHz	≈50 fs
OPA (Coherent, USA) - remnant SHG	≈405 nm	100 or 250 kHz	≈200 fs

Table 4.2: Laser system at the University of Zurich

The optical setup built for the 2PPE experiments is shown in Fig. 4.8. Ultrashort laser pulses were generated by a mode-locked Ti:Sapphire oscillator [94] (MIRA, Coherent, USA). The oscillator operated at the centre wavelength of 810 nm, a repetition rate of 80 MHz and the pulse duration was ≈ 50 fs. These laser pulses were amplified by the chirped pulse amplification (CPA) technique [95] in a regenerative amplifier (RegA, Coherent, USA). Prior to the amplification, the pulses were first stretched in time by 3 to 5 orders of magnitude in order to decrease the intensity.

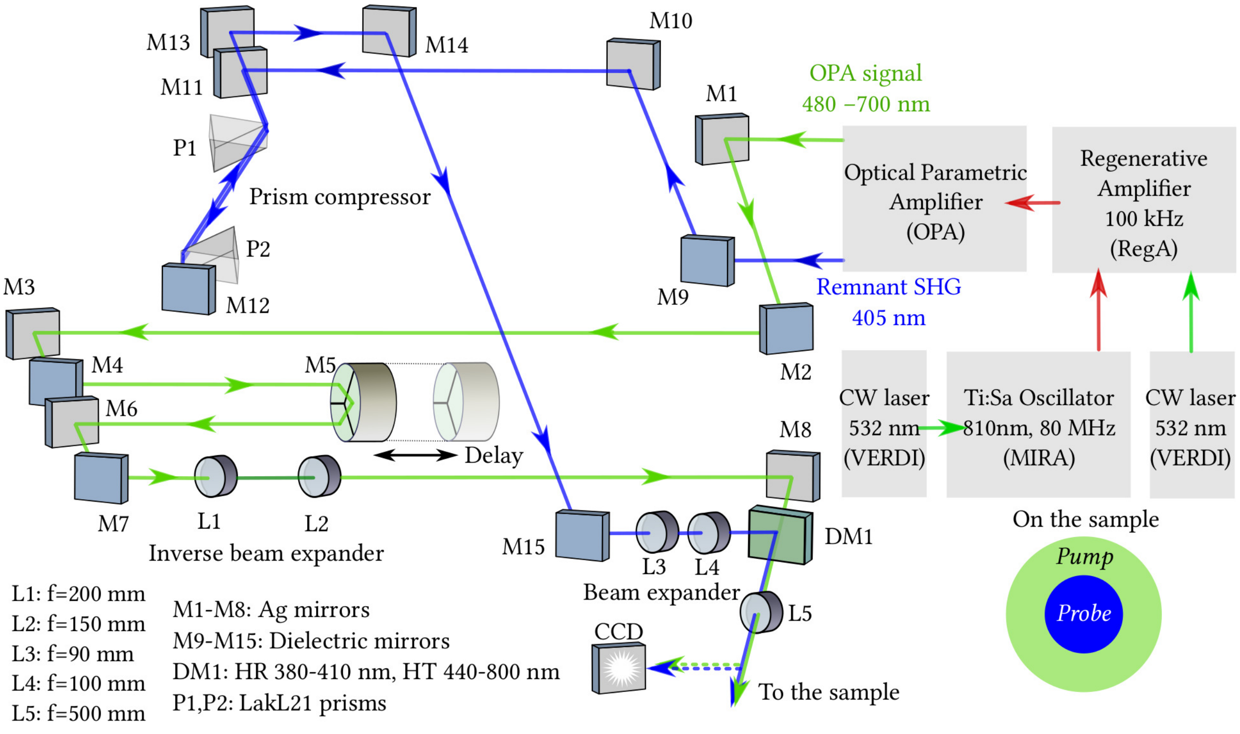


Figure 4.8: Laser setup for the 2PPE pump-probe measurements

The amplified pulses were then compressed back to almost initial pulse duration. Stretching and compression were realised with gratings, because refractive optics could be destroyed due to high fluences of the laser pulses. RegA was operating at the repetition rate of 100 kHz to ensure maximal energy per pulse. Amplified laser pulses from RegA were then used as an input for an OPA, an optical device that emits light at variable wavelengths by the process of optical parametric amplification [96]. This property of the OPA was particularly important as it allowed us to tune the excitation wavelength in a broad range. Fig. 4.9 shows the optical scheme of the OPA.

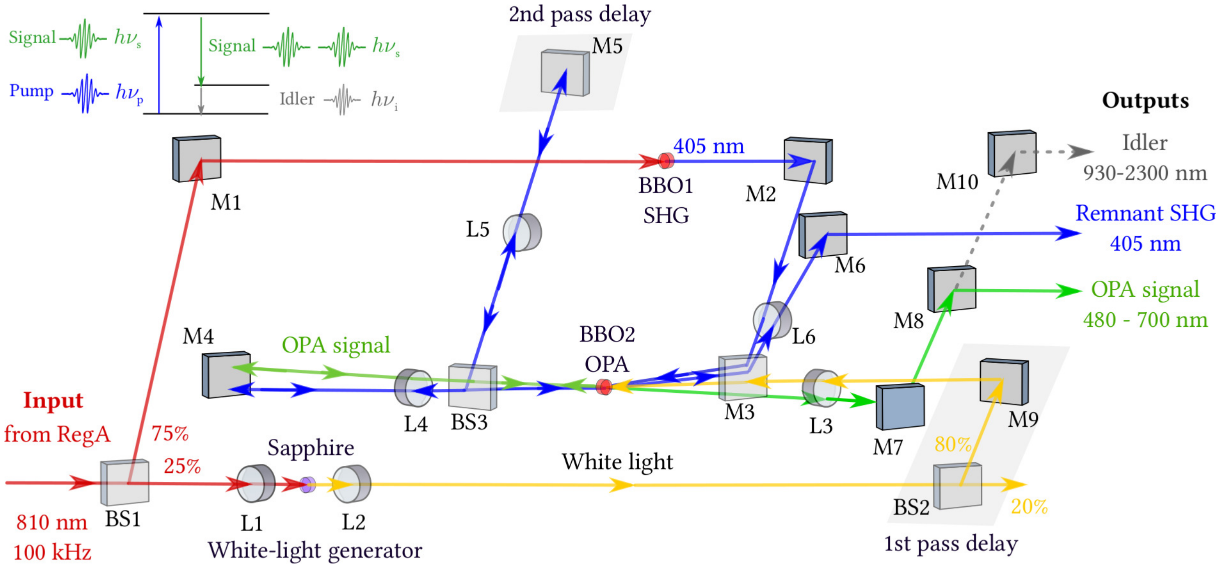


Figure 4.9: Optical scheme of the Optical Parametric Amplifier

75% of the light input was used for the second harmonic generation (SHG) [97], i.e. doubling of the input light frequency by means of a nonlinear process in a β−barium borate (BBO) crystal. Another 25% were used for the generation of a white-light continuum in a sapphire crystal. The white light was focused together with the SHG light into a second BBO crystal for the optical parametric amplification [96].

BBO is a nonlinear crystal, in which the polarisation of light quadratically depends on the electric field. This property allows for a conversion of the light with frequency ω_p (pump) into light with frequencies ω_s (signal) and ω_i (idler) satisfying the relation $\omega_s + \omega_i = \omega_p$, provided the corresponding light wave vectors satisfy the so-called phase-matching condition $\boldsymbol{K}_s + \boldsymbol{K}_i = \boldsymbol{K}_p$ [96]. The SHG light $\lambda = 405$ nm was used as the pump, while a part of the white-light continuum satisfying the phase matching condition was the signal. As a result of the optical parametric amplification process, this part of the white light was amplified and the OPA signal with pulse durations on the order of 50 fs was generated.[2] The OPA signal was used as pump in the 2PPE experiment. The remnant SHG light 405 nm was used as probe pulse.

Pump-probe experiments require spatial and temporal overlap of pump and probe laser pulses. Thus, the optical paths for pump and probe light (shown in green and blue in Fig. 4.8) were set to be equal. The pump beam was guided via a mechanical delay stage. The probe beam was slightly expanded using a Keplerian type of beam expander. The pump beam was shrinked in its diameter with an inverse Keplerian type of beam expander. Both beams were recombined on a dichroic mirror and focused onto the sample with a convex lens (f = 500 mm). The size of the pump beam on the sample exceeded the size of the probe beam approximately by a factor of 2 and was on the order of 150 μm.

4.5.2 *Pulse compression*

The remnant SHG pulses (405 nm) exiting the OPA were broadened due to their dispersion in the optical elements of the OPA and had to be compressed for the 2PEE experiment. The signal pulses were rather short in their duration due to a nonlinear nature of the parametric amplification process and did not require compression. For the compression of 405 nm laser pulses we constructed a pulse compressor consisting of a pair of LakL21 prisms [99]. The compression with prisms is based on the introduction of a negative group velocity dispersion (GVD) and group delay dispersion (GDD): the rear of the up-chirped pulse (the lower frequencies) propagate slower than the front (the higher frequencies) and the pulse is compressed in time. The ultimate pulse duration is then set by the width of the pulse spectrum, which remains unchanged upon linear pulse propagation.

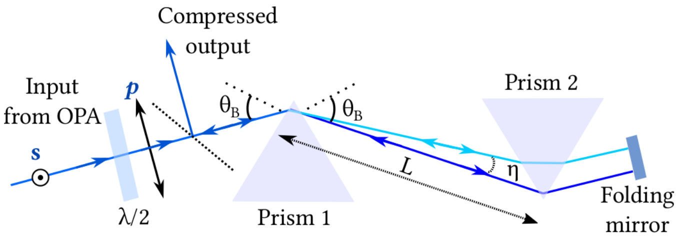


Figure 4.10: Prism compressor

The prism compressor is schematically shown in Fig. 4.10. The s-polarised laser beam exiting the OPA was rotated with a $\lambda/2$ plate. The p-polarised beam entered the first prism under the Brewster angle $\theta_B = \arctan(n)$, where n is the refractive index of the prism material, to reduce losses for p-polarised light. The apex angle of the prism was cut in such a way that the base angles were equal to the Brewster angle. This resulted in a symmetric beam path so that the beam exited the prisms also under the Brewster angle θ_B. The symmetric beam path with the equal incidence and exit angles allowed for a very elegant alignment of the first prism with respect to the incoming laser beam. The prism was rotated around the geometrical centre of its base and the deviation of the exit beam from the incidence beam direction was minimised. The entrance face of the second prism was placed parallel to the exit face of the first prism. The second prism collimated the beam that became angularly dispersed after the first prism. After the second prism, the different frequency components were separated in space due to the angular dispersion introduced by the first prism. To bring them back, a folding mirror was placed at the symmetry plane position, so that the the beam propagated back through the prism pair reversing the spatial chirp and doubling the temporal dispersion.[3]

As it was shown by Fork et al. [99], only the optical path $P = 2L\cos\eta(\lambda)$ contributes to the dispersion in this setup and the GDD thereby becomes equal to [100]

$$\mathrm{GDD}_{\mathrm{prism}} = \frac{\lambda^3}{2\pi c^2}\frac{d^2P(\lambda)}{d\lambda^2}, \tag{4.3}$$

---

[2] Shorter signal pulse durations up to 11 fs can be achieved using OPA in a non-collinear setting [98].

[3] Alternatively, one could place two additional prisms that mirror the disposition of the first two.

where L is the geometrical distance between the apexes of the prisms. Following the derivation of Fork et al. [99], and assuming that the product $L \cos \eta(\lambda)$ is approximately equal to twice the beam diameter at $1/e^2$ peak intensity[4], $2D_{1/\mathrm{e}^2}$, the GDD can be roughly estimated as

$$\mathrm{GDD}_{\mathrm{prism}} \approx 8 \cdot \frac{\lambda^3}{2\pi c^2} \cdot \left[-L \cdot \left(\frac{dn}{d\lambda} \right)^2 + \frac{d^2 n}{d\lambda^2} \cdot D_{1/\mathrm{e}^2} \right]. \tag{4.4}$$

The first term is always negative. Hence, by varying the prism separation L, the amount of the negative GDD can be roughly tuned. The second term is always positive and is attributed to the propagation of light through the prism material. By translating the second prism perpendicular to the beam path inside the prism, the amount negative GDD can be finely tuned.

4.6 LASER SETUP AT ARTEMIS

The ARTEMIS beamline [68] is located at the Central Laser Facility of the Rutherford Appleton Laboratory in the United Kingdom. The measurements at ARTEMIS aimed at studying the dynamics of the lowest unoccupied molecular orbital of the second layer of the pentacene bilayer on Ag(110). Upon resonant excitation of electrons in the pentacene second layer from HOMO to LUMO, they were to be probed with photons of high energy. Ideally, the photon energy of the probe light should be on the order of 25-35 eV, so that the photoelectrons with momenta between 1...2 Å^{-1}, containing most of the information about spatial distribution of orbitals, can be detected. For this purpose, the ARTEMIS beamline provides a unique combination of a laser source, tuneable from the infrared to ultraviolet, and ultrafast extreme ultraviolet (XUV) pulses produced through high harmonic generation (HHG) [101, 102].

The core device of the beamline is a 12 mJ, 30 fs, 1 kHz Ti:Sapphire chirped-pulse amplified laser system operating at 785 nm (RedDragon from KMLabs, USA). The laser system is carrier-envelope phase (CEP) stabilised. It is equipped with two grating compressors, allowing for the optimisation of pulse durations of pump and probe laser beams independently. A part of the output of the Ti:Sapphire laser system can be split and spectrally broadened in a gas-filled hollow fibre [103]. By focusing 1...2 mJ at 800 nm into the fibre and then recompressing the pulses with chirped mirrors, 0.5 mJ pulses as short as 7 fs can be produced.

Tuneable pulses in the spectral range from 230 nm to 20 μm are provided by an optical parametric amplifier (HE-Topas from Light Conversion, Lithuania). The OPA is pumped with up to 8 mJ of the output from the Ti:Sapphire laser system. At 1300 nm, pulse energies of up to 1 mJ with 40 fs pulse duration can be achieved. The availability of the OPA is important as it allows varying the pump wavelength in a broad spectral range and thus makes resonant excitation of HOMO-LUMO transitions possible.

XUV pulses in the spectral range of 10-100 nm (10-100 eV) are produced through the HHG [101, 102] process in a gas target by focusing the laser pulses from the Ti:Sapphire laser system to intensities of 10^{14} W·cm^{-2}. The XUV pulses are synchronised to the Ti:Sapphire laser pulses with sub-fs resolution. For the purposes of orbital tomography, the XUV radiation should be monochromatised. At ARTEMIS, this is done with a monochromator [104] consisting of two gold-coated toroidal mirrors and four interchangeable plane diffraction gratings to cover two spectral ranges (12-30 nm and 30-90 nm), providing either high resolving power or shorter pulse duration. The light is incident nearly parallel to the plane of the rulings to limit the dispersion of the pulses and maintain the resolving power and efficiency. Using the XUV monochromator instead of multilayer mirrors (i) allows for a simple tuning of the photon energy and resolution over the whole range of the HHG spectrum, and (ii) excludes the need for a metal filter, which improves the transmission through the beamline. At 32.5 eV photon energy (21st harmonic), XUV radiation has a flux of 1.6·10^{10} photons/second, corresponding to 80 pJ/pulse at 1 kHz and a monochromator bandwidth of 900 meV. At the cost of lower photon flux, a bandwidth of 390 meV can be achieved. For 28 fs IR driving pulses, the XUV pulse duration below 30 fs is achievable.

The XUV pulses are recombined with the pulses from the OPA inside a surface science end station for pump-probe experiments. The end-station is equipped with a sputter gun, e-beam heating, LEED device, He I source, a five-axis manipulator and a hemispherical electron energy analyser (SPECS PHOIBOS 100, Germany).

---

[4] Following the discussion of Fork et al. [99], the assume that the beam passes a distance equal to at least its diameter inside the apex of the first prism, in addition to the path through the second prism

Part II

RESULTS

OPTICAL ANALOGUE TO ORBITAL TOMOGRAPHY 5

The goal of an optical analogue experiment was to mimic the angle-resolved photoemission from an ordered molecular ensemble and demonstrate the viability of the iterative phase retrieval algorithms used in optical coherent diffraction imaging (CDI) for the reconstruction of molecular wave functions from photoemission data. The major part of this chapter was published as "P. Kliuiev, T. Latychevskaia, J. Osterwalder, M. Hengsberger and L. Castiglioni. Application of iterative phase-retrieval algorithms to ARPES orbital tomography. *New J. Phys.* **18** (2016) 093041" [PK3].

5.1 COHERENT DIFFRACTION IMAGING OF MICROSTRUCTURES

5.1.1 *Samples*

An important property of a sample for optical CDI [8] is that the sample distribution must be confined within a certain area. This ensures that the unknown sample distribution is surrounded by a medium with a known transmission function, thus mathematically providing the sample support, i.e. a set of all points that belong to the object [47]. Experimentally, such a sample support can be created by placing a sample inside a small aperture in an opaque screen. The ratio between the area of the aperture and the reconstructed object area is then the oversampling ratio which needs to exceed 2 [9, 10]. Another way to achieve known object support is to have the object in an opaque medium, as for example a transparent pattern in a non-transparent film. In this work, we employed the second type of object support.

The sample was manufactured by means of photolithography[1] and consisted of several different micrometer-sized structures (henceforth referred to as *microstructures*). The microstructures were patterned in a 105 nm-thick Cr film deposited on a 1.7 mm-thick fused silica substrate, thus providing transparent objects in a non-transparent medium. The main challenge in the sample design was a selection of sufficiently thick metal layer that is fully opaque to the laser light, but still is not too thick for the photolithographic patterning of the microstructures. By measuring the power transmitted through the Cr film at the wavelength of 532 nm, it was determined that only less then 1% of the incident light was transmitted through 105 nm of the metal, thus providing us with more than 99% opaque medium surrounding the microstructures. The individual microstructures had an identical shape but different sizes and were separated from one another by several millimetres to avoid interference between the neighbouring objects during the acquisition of the diffraction patterns.

Apart from a purely mathematical analogy in terms of the phase problem, the size of the microstructures was selected in such a way that the ratio between the microstructure length (e.g., 15 μm) and the employed laser wavelength (0.532 μm) was comparable to the ratio between the length of pentacene molecule ($\approx$ 1.5 nm) and de Broglie wavelength of the electrons ($\approx$ 0.17 nm) at photon energies typically used in ARPES experiments (50 eV).

The microstructures were manufactured in three different sizes. Sample 1: 30×12 μm^2, sample 2: 14.8×6 μm^2 and sample 3: 3×1.3 μm^2. The locations of the individual microstructures on the photomask were determined with an optical microscope Leica VZ 700 C. The bright- and dark images are shown in Fig. 5.1. Each type of the microstructure was then imaged with the scanning electron microscope[2] (SEM) ZEISS Supra 50 VP to discern the smallest details not visible in the optical microscope. Interestingly: as the dimensions of the smallest microstructure (sample 3) were close to the resolution limit of optical microscopy, its shape and size could be precisely determined only with the SEM. The samples 1 and 2 were produced with sharp features as it was originally designed. In case of the sample 3, it was not possible to manufacture the features finer than 400±100 nm due to technological limitations of the manufacturing process.

---

[1] The sample was manufactured by Compugraphics Jena GmbH (Germany) on behalf and using the final design of the ETH First Lab (Zurich, Switzerland)

[2] SEM imaging was performed with equipment maintained by the Centre for Microscopy and Image Analysis (University of Zurich).

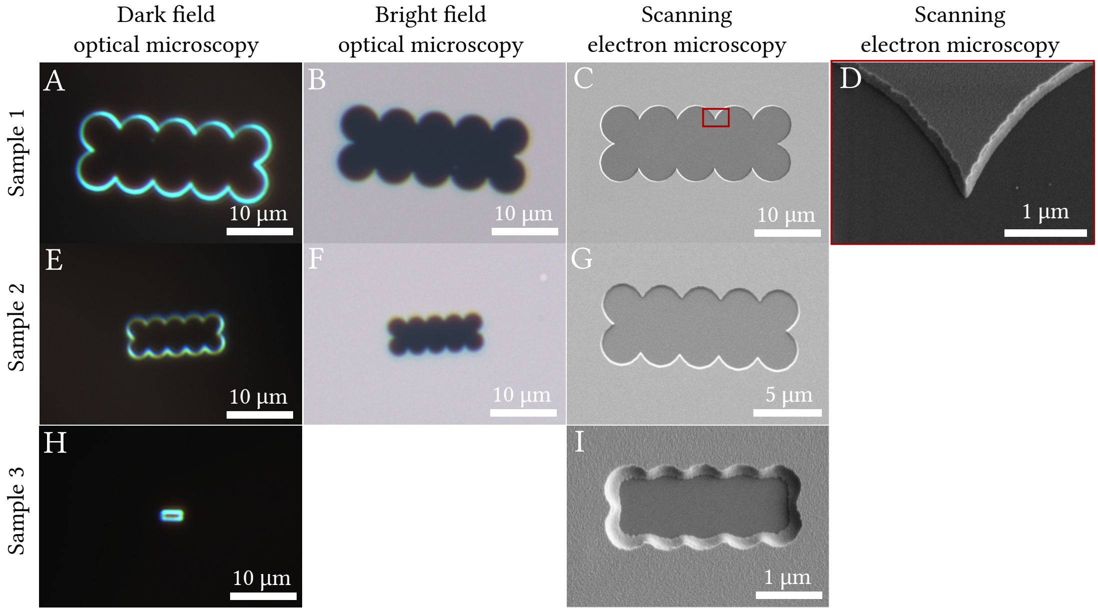


Figure 5.1: Optical and SEM images of the microstructures employed in optical CDI. (A-D) – Sample 1. (E-G) – Sample 2. (H,I) – Sample 3. (A,E,H) – Dark field optical microscopy images. (B,F) – Bright field optical microscopy images. (C,D,G,I) – SEM images. (D) – Zoom into an area with the sharpest features in Sample 1 marked by a red box in (C). (C,G): Adapted from [PK3].

5.1.2 *Setup for optical CDI*

The experimental setup for optical CDI is shown in Figure 5.2. A continuous wave (CW) Nd:YAG laser emitting at 532 nm was employed as a light source. The laser beam profile had a Gaussian distribution as shown in the inset. For CDI experiments, the laser beam is typically spatially filtered and its profile is expanded in size using a telescopic beam expander [105], which ensures that the intensity distribution in the object plane is constant. In our experiment, we employed the laser beam without expansion, since the intensity variations on the length-scale of the microstructures were negligible. The microstructure was illuminated from the side of the Cr film. The far field distribution of the diffracted light was imaged onto a semitransparent screen located at the distance of $z_s = 22.5$ cm from the sample plane. The diffraction pattern was recorded with a 10-bit CCD camera (Hamamatsu C4742-95) placed behind the screen as shown in Fig. 5.2.

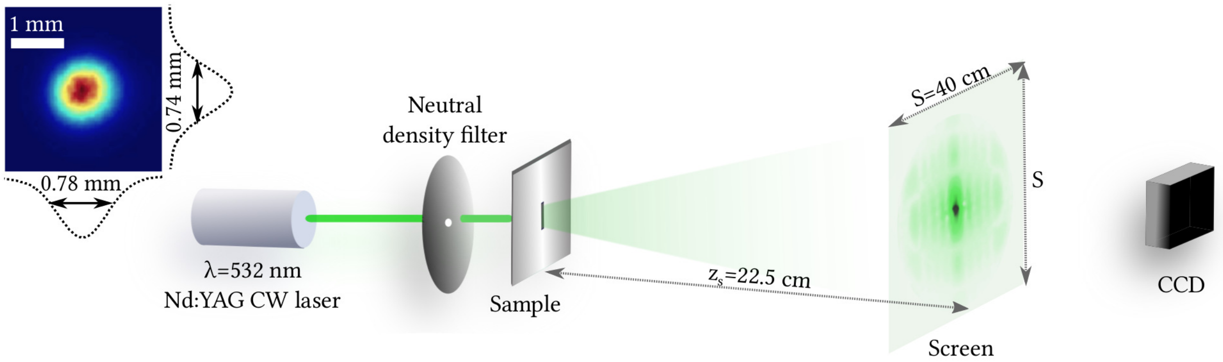


Figure 5.2: Experimental setup for optical CDI. The distance between the sample and the screen was set to 22.5 cm. The size of the part of the screen imaged with the CCD camera comprised 40×40 cm^2 and was sampled with 1000×1000 pixels. Adapted from [PK3].

In order to increase the dynamic range, we recorded several diffraction patterns at different exposures by using a rotatable neutral density filter with optical densities ranging from 0.0 to 4.0. The recorded images were then combined into one high-dynamic-range (HDR) image by a procedure proposed by Devebec [106].

5.1.3 *Some geometrical considerations*

As it was shown in chapter 2, in paraxial approximation, the coordinates in detector plane, $\tilde{x}$ and $\tilde{y}$, can be unambiguously related to the spatial frequencies of the light field, $q_x = \tilde{x}/\lambda z_s$ and $q_y = \tilde{y}/\lambda z_s$. If we interpret the diffraction as an elastic photon scattering process, then each pixel in the detector plane can be assigned a scattering vector $\Delta\, \boldsymbol{q}_{\text{sc}} = \frac{1}{2\pi}(\boldsymbol{K}_{\text{in}} - \boldsymbol{K}_{\text{out}})$, where $\boldsymbol{K}_{\text{in}}$ and $\boldsymbol{K}_{\text{out}}$ are the wave vectors of the incident and scattered photons: $|\boldsymbol{K}_{\text{in}}| = |\boldsymbol{K}_{\text{out}}| = K = \frac{2\pi}{\lambda}$.

If we consider an angle Θ for which the length of the scattering wave vector $|\Delta \boldsymbol{q}_{\text{sc}}| = \Delta q$ corresponds to the minimal possible sampling interval in reciprocal space limited by a detector pixel size, Δs, then $\Delta q \approx \Delta s/(\lambda z_s)$. Given the relation between the pixel sizes in real and reciprocal spaces, $\Delta r \Delta q = \frac{1}{N}$, where N is the linear number of pixels in the detector plane, we obtain for the pixel size in real space, Δr:

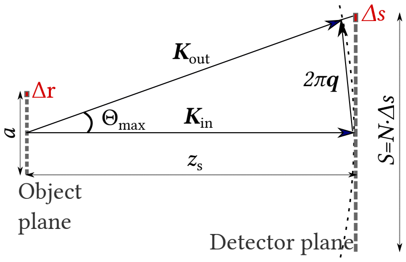


Figure 5.3: Scattering geometry in CDI

$$\Delta r = \frac{z_s \lambda}{N \Delta s}, \tag{5.1}$$

where z_s is the distance between object and detector planes. The linear oversampling ratio defined by Eq. 3.13, can then be computed as

$$\varnothing = \frac{z_s \lambda}{a \Delta s}. \tag{5.2}$$

5.1.4 *Reconstruction results*

Figure 5.4 shows the results of the reconstruction of the microstructure samples 1 and 2 only. The experimental diffraction patterns are shown in Fig. 5.4(A,D). They were sampled at $N \times N$ = 1000×1000 pixels and were each $S \times S$ = 40×40 cm^2 large. The size of the pixel in the detector plane was thereby Δs = 400 μm. The size of the pixel in the object plane was computed with Eq. 5.1 and comprised 300 nm. The linear oversampling ratios fulfilled the oversampling condition [9, 10]. For the samples 1 and 2 with the maximum lengths of a_1 = 30 μm and a_2 = 15 μm, they were equal to $\varnothing_1 \approx 10$ and $\varnothing_2 \approx 20$, respectively. It was not possible to record a diffraction pattern from sample 3 due to the absence of any optical signal in the far field, probably owing to the small size of the microstructure.

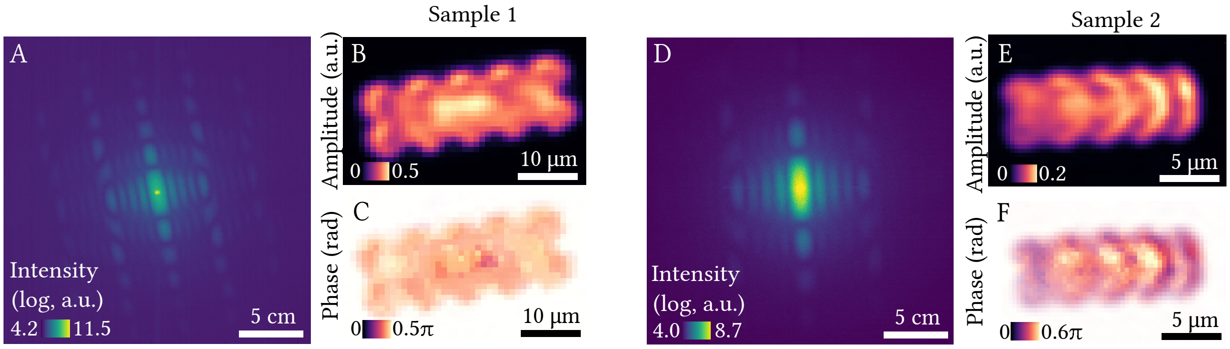


Figure 5.4: Reconstruction of the micrometer-sized objects. (A–C) – Sample 1. (D–F) – Sample 2. (A,D) – Experimental diffraction pattern intensities shown on logarithmic scale. (B,E) – Reconstructed amplitude distributions. (C,F) – Reconstructed phase distributions weighted with the corresponding amplitude values for illustration purposes. Adapted from [PK3].

Prior to application of the phase retrieval algorithms, the experimental diffraction patterns were preprocessed. (i) Each of the recorded 1000×1000 pixel images was centred. Centering of the experimental diffraction pattern was shown to have a strong effect on the quality of the reconstruction in CDI [107]. (ii) The noise of the CCD camera (average count rate of 50 counts) was subtracted from each pixel and (iii) the images were truncated to $N' \times N'$ = 500×500 pixels around their centres because of the low signal-to-noise ratio in the peripheral parts. This resulted in a new size of the pixel in the object domain

$\Delta r' = 600$ nm. (iv) The central part of each diffraction pattern was dominated by an intense laser spot due to the partial transparency of the chromium film. Thereby, the pixel values exceeding the thresholds of $1.5{\cdot}10^5$ counts (sample 1) and $6{\cdot}10^3$ counts (sample 2) were defined as missing and their values were updated in the course of the reconstruction by using the corresponding pixel values of the calculated amplitudes in the detector plane [12, 58]. We note that no correction of the diffraction patterns due to their projection onto the flat screen was done.

The square root of the processed diffraction patterns was fed into the phase retrieval procedure described in chapter 3. We found that 10 alternating cycles of the PC-HIO and ER algorithms, each followed by an update of the support, were enough to achieve a stable reconstruction. Further increase in the number of the reconstruction cycles was not necessary since it did not improve the quality of the reconstructed object distribution. At the end of 10 cycles, each reconstruction was stabilised by 100 iterations of the ER algorithm [58]. In total, we performed 1000 independent reconstructions of each microstructure by employing a random phase distribution for each reconstruction run. Eventually, the 50 reconstructions with the smallest error E_{R}, as defined by Eq. 3.3, were selected and averaged [54, 58–61] and are shown in Fig. 5.4(B,C,E,F). The reconstructed amplitudes correctly reproduce the shape and dimension of the microstructures. Furthermore, as it was expected for a purely transmitting object illuminated by a Gaussian beam with an almost planar wavefront at the object site, the phase distributions turned out to be almost constant. The lower quality of the reconstructed amplitude of the sample 2 in Fig. 5.4(E) can be attributed to the lower signal-to-noise ratio in the respective diffraction pattern.

5.2 ARPES ORBITAL TOMOGRAPHY OF PENTACENE ON AG(110)

By employing the analogy between the phase problem in optical CDI and orbital tomography, we then applied the same phase retrieval algorithm to a set of the ARPES data.

5.2.1 *Sample and experimental setup*

As a model system, we used a well-ordered sub-monolayer of pentacene molecules adsorbed on an Ag(110) substrate. Pentacene ARPES data were acquired during a beamtime of A. Schöll and coworkers (University of Würzburg) at the NanoESCA beamline at Elettra synchrotron (Trieste, Italy) and provided to us for validation of our phase retrieval routine. The crystal was prepared according to standard procedures and pentacene molecules were deposited from a home-built Knudsen cell. PADs were recorded at the binding energy of the pentacene lowest unoccupied molecular orbital (LUMO) with p-polarised light and photon energy of 50 eV using the photoemission electron microscope [71, 108]. The setup of the PEEM and the experimental geometry were analogous to those shown in Fig. 4.2. The microscope was operated in the momentum mode and allowed for the detection of electrons within the field of view ± 2 Å^{-1} without any sample rotation. The PAD was integrated over a 200 meV energy window, which is on the order of the electron analyser resolution and of the full-width at half-maximum of the pentacene LUMO at the binding energy of 0.1 eV.

5.2.2 *Reconstruction results*

Fig. 5.5 shows the results of the reconstruction of the pentacene LUMO. The experimental PAD pattern is shown in Fig. 5.5(A). The linear oversampling ratio in the ARPES experiment is

$$\O = \frac{2\pi}{a\Delta k}, \tag{5.3}$$

where Eq. 3.13 and the relation between the pixel size in the object space, Δr, and reciprocal space, Δk, $\Delta r \Delta k = \frac{2\pi}{N}$, were taken into account. Given the resolution in reciprocal space of $\Delta k \approx 0.01$ Å^{-1} and the length of the pentacene molecule $a \approx 15$ Å, the linear oversampling ratio was $\O \approx 42$ and thus fulfilled the oversampling condition [9, 10].

The experimental PAD pattern was pre-processed following the steps similar to those applied prior to the reconstruction of the micrometer-sized objects. (i) The image was centred and (ii) the quasi-constant noise of the CCD camera (average count rate of 50 counts) was subtracted from each pixel. (iii) To ensure a sufficient number of pixels allocated per unit length of the molecule, we zero-padded the experimental PAD pattern to 2000×2000 pixels around its centre. (iv) The square root of the processed data were

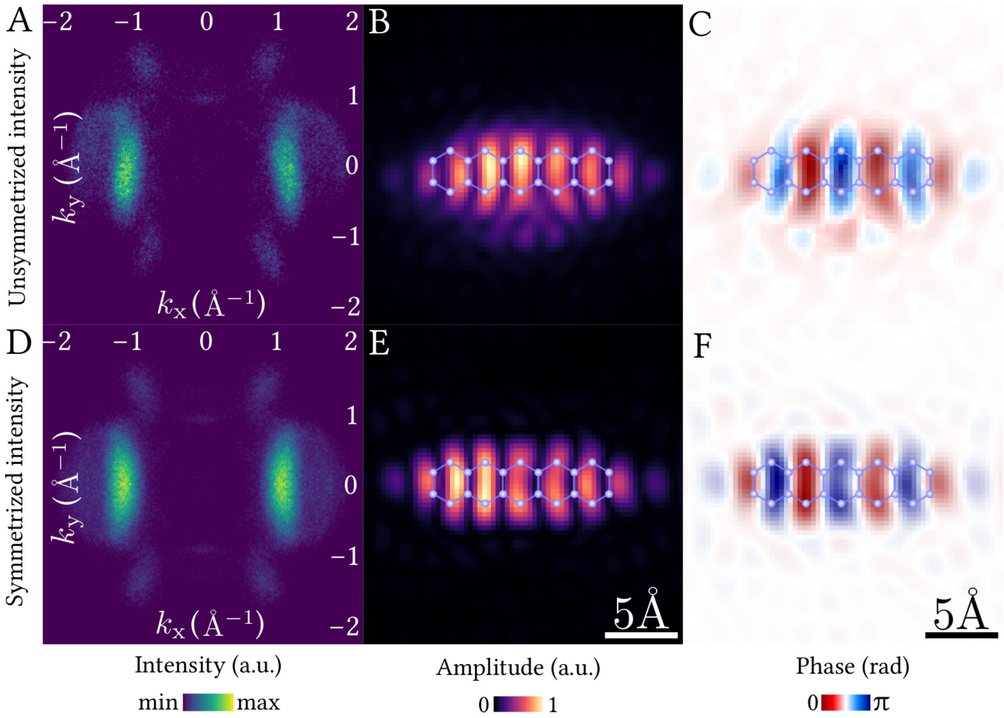


Figure 5.5: Reconstruction of the pentacene LUMO. (A,D) – PAD distributions. (B,E) – Reconstructed amplitude distribution in real space. (C,F) – Reconstructed phase distributions in real space weighted with the corresponding amplitude values for illustration purposes. (A–C) – Reconstruction from unsymmetrised data (A). (D–F) – Reconstructions from symmetrised data (D). Adapted from [PK3].

then fed into the algorithm with the same parameters as used for the reconstruction of the micrometer-sized structures. Varying these parameters did not lead to any substantial improvements in the quality of the reconstruction. In total, we performed 1000 reconstructions of the pentacene LUMO. About 56% of the reconstructed objects $g(x, y)$ were reconstructed together with their conjugate (twin) images $g^*(-x, -y)$ [53]. The identification of the twin images could be automated by a procedure proposed by Fienup [53], but here they were easily identified by visual inspection and discarded. From the remaining reconstructions, 50 with the smallest error E_R, given by Eq. 3.3, were selected and averaged [54, 58–61]. The reconstructed amplitude and phase of the pentacene LUMO are shown in Fig. 5.5 (B,C) together with the overlaid carbon frame of the molecule for comparison.

It should be noted that in this particular case we did not perform a normalisation of the ARPES intensity by the angle-dependent factor $|\boldsymbol{A} \cdot \boldsymbol{k}_f|^2$ nor did we enforce any symmetry constraints in the course of the reconstruction onto the amplitude and phase shown in Fig. 5.5(B,C). The object distribution was let to freely evolve until the stable solution was reached, which makes the utilised algorithm independent of any symmetry properties imposed onto the object under reconstruction. Furthermore, we note that the recorded PAD pattern shown in Fig. 5.5(A) contains features coming from the Ag(110) substrate (mostly at high momenta), but they do not seem to have a profound effect on the results of the reconstruction. By comparing our results with the literature, we find that the phase distribution weighed with the correspondent amplitude values as well as the shape of the orbital correctly reproduce the DFT calculations [3, 109] as well as the data reconstructed by Lüftner et. al [3].

Finally, in order to assess the robustness of the algorithm in terms of the quality of reconstruction from the unsymmetrised PAD pattern, we made use of the symmetry properties of the pentacene LUMO amplitude and phase and symmetrised the PAD shown in Fig. 5.5(A) around its centre. The symmetrical version is shown in Fig. 5.5 (D). In optics, the far field diffraction pattern is symmetric only in two cases: either due to the real-valued nature of an object or in case of an even complex-valued object distribution with an even amplitude and an even phase. In the latter case, the Fourier transform of the even complex-valued function is an even function as well and the far field intensity distribution is therefore symmetric. In the case of the complex-valued wave function of the pentacene LUMO, the symmetrisation is justified purely due to the symmetry of the LUMO amplitude and the phase as it is known from the DFT calculations [3, 109]. The symmetrised PAD was pre-processed using the same procedure as described above and the results of the reconstruction are shown in Fig. 5.5(E,F). Qualitatively, the reconstructions from the unsymmetrised PAD are as good as the reconstructions from the symmetrised data set, except for some minor differences in the shapes of the lobes due to the intrinsic asymmetry of the PAD in Fig. 5.5(A). This agreement further proves the robustness of the employed algorithm for the reconstruction of molecular orbitals with arbitrary symmetry properties.

6

DATA PROCESSING AND IMAGE FORMATION IN ORBITAL TOMOGRAPHY

In orbital tomography, the molecules are adsorbed on a single crystal substrate, which ensures that the molecules are well ordered. It is also required that the interaction between molecular species and the substrate is weak [1–3]. This allows to reconstruct the orbitals of quasi "free-standing" molecules and then directly compare them with the gas phase simulations [1–3]. However, the PADs, recorded at binding energies of molecular states, might contain not only features due to photoemission from "pure" molecular states, but also some signal from the substrate that might be present at the same binding energies [110]. Since such substrate features do not originate from the molecular states to be probed, they should not be present in the PAD data for orbital reconstruction. This issue can be solved either (i) by choosing a substrate with a low density of states at the respective binding energies of molecular states [1–5], which however is not always possible or desirable, or (ii) by applying a fitting procedure [110]. Unfortunately, these fitting algorithms may not eliminate all spurious features completely [110]. In addition, they require acquisition of complete PADs throughout a broad range of binding energies at a sufficiently dense sampling rate. And since the phase retrieval algorithms perform better when the data were acquired with higher statistics, the acquisition time may therefore increase substantially. To solve this issue, we suggest using a signal restoration procedure consisting of registration of signal and background data via maximisation of mutual information [14, 111]. We show that this procedure effectively removes the principal background features. As a result, instead of recording data at a broad range of binding energies with high statistics and applying the fitting procedure, we suggest performing a fast survey followed by acquisition of PADs with high statistics only at binding energies of the identified molecular states.

By applying this signal restoration procedure to the ARPES data recorded from the highest occupied molecular orbital (HOMO) of pentacene on Ag(110) and employing the phase retrieval algorithms [5, 11–13, 49], we were able to reconstruct both the amplitude and phase distribution of the orbital without any symmetrisation of the data or prior information about parity or shape of the wave function. The reconstruction was done using an improved version of our reconstruction procedure [5], by adding an additional refinement step, allowing for the precise determination of the geometrical centres of the reconstructed orbital distributions. Most importantly, by making a comparison with photoelectron holography [18], we interpret two-dimensional (2D) reconstructed orbital distributions as a superposition of "in-focus" and "out-of-focus" contributions from the original three-dimensional (3D) orbitals. This remarkable result means that the full 3D orbital distributions can be reconstructed solely from a single set of 2D ARPES data, provided the axial resolution is high enough.

6.1 SAMPLE

We employed ARPES data from the highest occupied molecular orbital of pentacene on Ag(110) recorded using p-polarised light of 40 eV photon energy with the PEEM [71, 72]. The data were acquired during a beamtime of Schöll and co-workers (University of Würzburg) at the NanoESCA beamline at Elettra synchrotron and has been provided to us for validation of our algorithms. The experimental geometry was identical to that shown in chapter 4, Fig. 4.2. Molecules were deposited using standard procedures listed in chapter 4 .

6.2 DATA PROCESSING

6.2.1 *Registration of experimental data*

In order to obtain sufficiently high signal-to-noise ratio of the pentacene valence state PAD, Schöll et al. acquired data in a 200 meV energy window, which is of the order of the electron analyser resolution and of the full-width at half-maximum of the pentacene HOMO at the binding energy of 1.2 eV. The data were taken in steps of 20 meV with the dwell time of 3 s per image. In total, Schöll et al. did 5 independent recordings of the data in this energy window and summed up the resulting 50 PADs. Because organic molecules suffer from radiation damage under intense UV and x-ray irradiation, the PADs were acquired with short acquisition times and raster scanning of the sample. This caused image drift and, as a result, the normal emission direction in the individual PADs was not aligned to a common pixel on the CCD. As a result, prior to summation, we had to perform numerical registration of the data, i.e. the overlaying of all 50 PAD images acquired at different times with the purpose of their geometrical alignment [15].

The PADs, $I = I(\boldsymbol{k}_{\mathrm{f}\|}, E_{\mathrm{kin}})$, were sampled at $N \times N$=540×540 pixels and the mutual translation $\boldsymbol{\Delta}$ was only in the order of 8 pixels both in horizontal and vertical directions. Therefore, for the registration, we chose the method of intensity interpolation based on the cross-correlation [114]. This method outperforms all other relevant registration methods [115] in alignment of a real-valued image sequence contaminated with noise [114]. The amount of shift was found as follows:

1. The experimental PADs were up-sampled by a factor of s=2 to $N' \times N'$=1080×1080 pixels by bicubic interpolation.
2. Then they were normalised as $I'_{\mathrm{norm}}(u,v) = I'(u,v)/\sqrt{\sum_{u,v=0}^{N'-1} I'(u,v)}$, where u and v are the coordinates in the detector plane and the prime symbol denotes the up-sampling.
3. The mean value was subtracted from each pixel of the normalised images [116].
4. The position of the cross-correlation maximum of the first two up-sampled images, $\boldsymbol{\Delta}'$, delivered the relative shift between the experimental PAD images, $\boldsymbol{\Delta} = \boldsymbol{\Delta}'/s$, which was used to merge them by summation.
5. The merged image was then used as a signal to bring it into registry with the next image.
6. The process was repeated sequentially until the whole image sequence was aligned.

The PAD obtained upon registration of 50 data sets is shown in Fig. 6.1(A). The broad blobs in Fig. 6.1(A) are attributed to photoemission from the molecular state. The narrow sharp features crossing them originate from the sp-bands of the substrate and hence have to be eliminated prior to the application of the phase retrieval procedure. The corresponding PAD recorded from the bare Ag(110) substrate at 1.2 eV binding energy, 5 s acquisition time, is shown in Fig. 6.1(B).

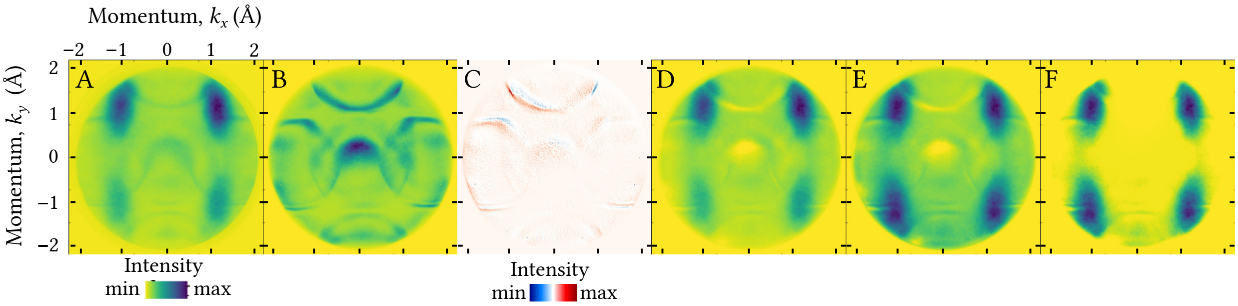


Figure 6.1: Processing of the pentacene HOMO PAD. (A) – Sum over 50 geometrically aligned PADs recorded from a sub-monolayer of pentacene molecules on Ag(110) in a 200 meV energy window at 1.2 eV binding energy, 3 s acquisition time per image. (B) – Background PAD acquired from the Ag(110) substrate at 1.2 eV binding energy, 5 s acquisition time. (C) – Difference between the background PAD in (B) and the same PAD, but registered with the signal PAD in (A) using maximisation of mutual information algorithm. (D) – The same signal PAD as in (A), but after the subtraction of the registered background. (E) – The same PAD as in (D), but normalised with the $|\boldsymbol{A} \cdot \boldsymbol{k}_{\mathrm{f}}|^2$ factor. (F) The same PAD as in (E), but after subtraction of the mean value from each pixel. Adapted from [PK2].

6.2.2 *Alignment of signal and background data via maximisation of mutual information*

The removal of substrate bands from the signal data cannot be done by direct subtraction of the two raw images due to slight shifts and distortions between them. Therefore, the background data shown in Fig. 6.1(B) has to be brought into spatial registry with the corresponding features of the signal data shown in Fig. 6.1(A) before subtracting the substrate contribution. For this purpose, instead of using the registration method based on the cross-correlation [114], which accounts only for the image translation, we employ a method of maximisation of mutual information (MI) [111, 121–124] based on the algorithm of Mattes [14, 125]. The algorithm was implemented in MATLAB. Here we briefly describe only the most important points of this algorithms necessary for its understanding.

Alignment by maximisation of mutual information is a very robust technique that does not require any pre-processing of the images or information about the nature of the imaging process [124]. Let us denote the signal and background images shown in Fig. 6.1(A, B) by $I_S(\boldsymbol{r})$ and $I_B(\boldsymbol{r})$, respectively and seek a geometrical transformation $\mathcal{G}$ between the two respective image coordinate systems. We assume that $I_S(\boldsymbol{r})$ and $I_B(\boldsymbol{r})$ are defined on continuous domains by interpolation between the corresponding grid points, as described later in the text. The transformation $\mathcal{G}$ is a function of transformation parameters μ that maps points $\boldsymbol{r}$ in the signal image to points $\mathcal{G}(\boldsymbol{r})$ in the background image, so that they have the same physical origin, or, in medical terms, the same "anatomy". We seek for a set of transformation parameters that minimises a negative of an image similarity function $\mathcal{C}$:

$$\hat{\mu} = -\arg\min_{\mu} \mathcal{C}. \tag{6.1}$$

$\mathcal{C}$ is a very general function that can be, for instance, a weighted average of the standard deviation of the pixels in the background image within each partition of the signal image, as it was proposed by Woods et al. [126]. While such a criterion may be effective in some applications, in general, the image similarity function should be non-heuristic in its nature. Thus, in 1995, Collignon et al. [121] proposed using mutual information, $\mathcal{M}(\rho_S;\rho_B)$, [127], as a similarity metric:

$$\mathcal{M}(\rho_S;\rho_B) = \mathcal{S}(\rho_S) - \mathcal{S}(\rho_S|\rho_B) = \sum_{\rho_S,\rho_B} p_{SB} \log \frac{p_{SB}}{p_S p_B}. \tag{6.2}$$

In Eq. 6.2, ρ_S and ρ_B are random variables, $\mathcal{S}(p_S)$ is the marginal entropy of ρ_S, $\mathcal{S}(\rho_S|\rho_B)$ is the corresponding conditional entropy, p_S and p_B are the marginal probability distributions of ρ_S and ρ_B, respectively, and p_{SB} is the joint probability distribution. Collignon arrived at the conclusion that due to properties of the mutual information, ρ_S and ρ_B can be replaced by the grey values I_S and I_B of the images to be registered, so that $\mathcal{M}(\rho_S;\rho_B)$ becomes $\mathcal{C}(I_S;I_B)$.

Due to the continuity of the transformation parameters $\hat{\mu}$ [121] and in order to facilitate the optimisation process [125], the mutual information function has to be explicitly differentiable. In turn, this requires the background image, all probability distributions and the transformation function to be differentiable as well. In practice, this can be achieved by interpolation of the corresponding pixel values between the grid points. In the algorithm of Mattes [14, 125], the continuous version of the discrete intensity distribution is represented by a B-spline basis:

$$I(\boldsymbol{r}) = \sum_i c_i \beta^{(3)}(\boldsymbol{r} - \boldsymbol{r}_i), \tag{6.3}$$

where $\boldsymbol{r}_i$ is a discrete set of coordinates defined on a Cartesian grid, $\boldsymbol{r} \in \mathbb{R}$ is a continuos coordinate on this grid, $\beta^{(3)}(\boldsymbol{r}) = \beta^{(3)}(x)\beta^{(3)}(y)$ is a separable convolution kernel and the coefficients c_i are computed from the corresponding background image samples using a recursive filtering algorithm [128]. The argument of the convolution kernel is the cubic B-spline

$$\beta^{(3)}(x) = \begin{cases} \frac{1}{6}(4 - 6x^2 + 3|x|^3), & 0 \le |x| < 1 \\ \frac{1}{6}(2 - |x|)^3, & 1 \le |x| < 2 \\ 0, & 2 \le |x|. \end{cases} \tag{6.4}$$

The probability distributions are estimated by computing the corresponding image histograms and forming continuous estimates with Parzen windows [129]. The histogram bins in the signal and background images are given by integer values κ and ι, respectively, and correspond to linearly scaled intensity values. The determination of the number of bins in histograms is a very controversial topic in image processing [130]. In this work, the number of bins was determined from the background image by using Scott's rule [131] with the skewness factor [132].

The smoothed joint probability distribution is computed as

$$p(\iota,\kappa|\mu) = c_0 \sum_{\boldsymbol{r}} \beta^{(0)}\left(\kappa - \frac{I_S(\boldsymbol{r}) - I_S^o}{\Delta b_S}\right) \beta^{(3)}\left(\iota - \frac{I_B(\mathcal{G}(\boldsymbol{r}|\hat{\mu})) - I_B^o}{\Delta b_B}\right), \tag{6.5}$$

where c_0 is a normalisation coefficient[1], $I_S(\boldsymbol{r})$ and $I_B(\mathcal{G}(\boldsymbol{r}|\hat{\mu}))$ are the samples of the signal and background images, respectively, I_S^o and f_L^o are the minimum intensity values, Δb_S and Δb_B are the intensity ranges of each bin. The summation is done over a subset of pixel locations in the reference image that contribute to the distribution. The smoothing is done by zero-order and cubic spline Parzen windows, denoted by $\beta^{(0)}$ and $\beta^{(3)}$, respectively, both of which satisfy the partition of unity constraint [129]. The marginal smoothed histogram of the background image is computed from the smoothed joint histogram by summation over all bins in the signal histogram:

$$p_B(\iota|\hat{\mu}) = \sum_{\kappa}(\iota,\kappa|\hat{\mu}). \tag{6.6}$$

The marginal smoothed histogram of the signal image is computed independently of the transformation parameters, since the B-spline Parzen window satisfies the partition of unity constraint:

$$p_S(\kappa) = c_0 \sum_{\boldsymbol{r}} \beta^{(0)} \left(\kappa - \frac{I_S(\boldsymbol{r}) - I_S^o}{\Delta b_S} \right). \tag{6.7}$$

The negative of mutual information given by Eq. 6.2 can then be rewritten as a continuous estimate:

$$\mathcal{M}(\hat{\mu}) = -\sum_{\iota}\sum_{\kappa} p(\iota,\kappa|\hat{\mu}) \log \frac{p(\iota,\kappa|\hat{\mu})}{p_B(\iota|\hat{\mu})p_S(\kappa)}. \tag{6.8}$$

Computation of the mutual information with Eq. 6.8 provides a measure of statistical dependence between the two images for a given geometrical transformation $\mathcal{G}$. The optimisation process itself, the search of the transformation function $\mathcal{G}$ that minimises the metrics, can in principle be driven by a number of different optimisation algorithms.

In this work, we used the one-plus-one evolutionary algorithm [133]. The transformation function $\mathcal{G}$ accounted for affine transformations only and was defined as follows:

$$\mathcal{G} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix}}_{\textit{Translation}} \cdot \underbrace{\begin{bmatrix} \cos\vartheta & \sin\vartheta & 0 \\ -\sin\vartheta & \cos\vartheta & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\textit{Rotation}} \cdot \underbrace{\begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\textit{Scale}} \cdot \underbrace{\begin{bmatrix} 1 & \mathrm{sh}_y & 0 \\ \mathrm{sh}_x & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\textit{Shear}}, \tag{6.9}$$

where t_x and t_y are the displacements, s_x and s_y are the scale factors, sh_x and sh_y are the shear factors along the x and y axes, respectively. The angle of rotation about the origin is ϑ. In the one-plus-one evolutionary algorithm, the set of transformation parameters $\hat{\mu} = \{t_x, t_y, s_x, s_y, \mathrm{sh}_x, \mathrm{sh}_y, \vartheta\}$ is called "population". The algorithm started with unperturbed parameters, i.e. all values of μ were set to 0. The first guess was generated randomly, using an initial search radius c_{init} around the initial position in the search space. This means, that in the first iteration, each parameter attained a new random value between 0 and c_{init}. In the second iteration, the parameters from the first iteration (parent) were perturbed, yielding a new distribution of the parameters (child). Among these parameters, only those yielding the lowest similarity metric were retained for the following optimisation step. The search radius for the next step was incremented, as set by the growth factor c_{grow}. The optimisation was done in $n = 4$ cycles, each consisting of 1000 iterations. The initial search radius was set to $c_{\text{init}} = 10^{-3}$, the growth factor to $c_{\text{grow}} = 1.01$. The number of histogram bins determined by the Scott's rule [131] was 127. In the first cycle, the linear number of pixels was set to $N/(2n)$ in each dimension and then increased at the beginning of each new cycle by a factor of 2, until in the last cycle it was again equal to $N = 540$, i.e. to the linear number of pixels in the signal and background images. This procedure allowed for a gradual refinement of the optimisation results, until the algorithm converged after 200-250 iterations in the last cycle. The difference between the background data after and before registration is shown in Fig. 6.1(C). The registered background was subtracted from the signal image and the resulting background free image, I_s, is shown in Fig. 6.1(D). It is seen already by a visual inspection that most of the background features become suppressed upon subtraction. The mutual information of the registered background and signal data was equal to 0.3565, given that the mutual information of the two identical sets of the background data equals to unity. The mutual information of the background and signal data aligned by the intensity interpolation method, in which only a translation was taken into account, was equal

---

[1] It ensures $\sum_{\iota,\kappa} p(\iota,\kappa) = 1$

to 0.1789. Thus, as a result of alignment with the maximisation of mutual information algorithm, in which the background image endured a set of affine transformations, the registered background data contained ≈ 2 times more information about the background features in the signal data, compared to image registration with the intensity interpolation.

6.2.3 *Normalisation with the $|\boldsymbol{A} \cdot \boldsymbol{k}_\mathrm{f}|^2$ factor*

After removal of the background features, the PAD intensity distribution was normalised. In the experimental geometry, shown in chapter 4, Fig. 4.2, the angle of incidence of the incoming light, $\alpha = 25°$, was kept fixed, while the photoelectrons were detected at a broad range of polar angles θ. Thus, the intensity distribution in the detector plane became modulated by the $|\boldsymbol{A} \cdot \boldsymbol{k}_\mathrm{f}|^2$ factor, as set by Eq. 1.21. To eliminate this modulation, we divided the PAD shown in Fig. 6.1 (D) by

$$|\boldsymbol{A} \cdot \boldsymbol{k}_\mathrm{f}|^2 = |A_x k_{\mathrm{f},x} + A_y k_{\mathrm{f},y} + A_z k_{\mathrm{f},z}|^2 = \left| A_x k_{\mathrm{f},x} + A_y k_{\mathrm{f},y} + A_z \sqrt{\frac{2m}{\hbar^2}(E_\mathrm{kin} + V_0) - k_{\mathrm{f},x}^2 - k_{\mathrm{f},y}^2} \right|^2, \quad (6.10)$$

where the electric field vector potential components $A_x = \sin\alpha \cdot \sin\varphi$, $A_y = \sin\alpha \cdot \cos\varphi$, $A_z = \cos\alpha$ and the polar angle of incidence, $\alpha = 25°$, were defined by the experimental geometry. In the experiment, the azimuthal angle φ was 5°. The mean inner potential, V_0, which typically varies between 5 to 9 eV for overlayers of common organic molecules [117, 118], was set to $V_0 = 7$ eV, as this value led to the best results in terms of the symmetry of the normalised data. The resulting PAD is shown in Fig. 6.1(E). After the division of the experimental PAD by the $|\boldsymbol{A} \cdot \boldsymbol{k}_\mathrm{f}|^2$ factor, the mean intensity value was subtracted from each pixel of the normalised PAD distribution and the values of all negative pixels were set to 0. The resulting PAD, shown in Fig. 6.1(F), was used as the sole input for the iterative phase retrieval procedure.

6.3 ITERATIVE RECONSTRUCTION OF THE PENTACENE ORBITAL

In the course of image reconstruction with phase retrieval algorithms, one often encounters the problem of a translated support distribution. Because an orbital distribution $\psi_\mathrm{rec}(l, m)$ and its duplicate shifted by (l_0, m_0) pixels, $\psi_\mathrm{rec}(l - l_0, m - m_0)$, have the same amplitude of the Fourier transform, the location of the object support, obtained with the shrinkwrap algorithm, is arbitrary and partially reconstructed orbital distributions are often not aligned with the support constraint [53]. As a consequence, one faces difficulties with determining the geometrical centre of the reconstructed object distribution, even after averaging over multiple reconstructions. To illustrate this problem, let us consider the results of a reconstruction done using the processed ARPES data of pentacene shown in Fig. 6.1(F).

6.3.1 *Reconstruction of the pentacene HOMO using conventional phase retrieval procedure*

The reconstruction was done using alternating cycles of the phase-constrained [13] hybrid input-output [11] (PC-HIO) and the error reduction [11] (ER) algorithms. The object support was obtained using the shrinkwrap algorithm [12]. For that, the initial estimate of the object support was obtained by computing the inverse Fourier transform of the processed experimental PAD data, I_s, convolving it with a Gaussian function (standard deviation $\sigma_0 = 3$ pixels), thresholding at 9% of its maximum and setting the pixel values below the threshold to zero. In the last iteration of each ER cycle, the output object distribution was used to update the object support by convolving it with a Gaussian function and setting a threshold at 19% of its maximum. The width of the Gaussian was initially set to $\sigma_0 = 2.5$ pixels and was reduced by 1% at every support update. In total, we performed 1000 independent reconstruction rounds with different initial random phase distributions. Each reconstruction round consisted of 5 alternating cycles of 10 iterations of the PC-HIO algorithm, 5 iterations of the ER algorithm and an update of the support. At the end of five cycles, each reconstruction was stabilised by 100 iterations of the ER algorithm [5, 58]. We selected only 10% of the object distributions having the lowest error metric in the reciprocal space [11, 49, 53] and averaged them [5, 58–60].

The results of the reconstruction are shown in Fig. 6.2(A,B). The spatial resolution in the object domain was estimated as $\Delta r = \frac{2\pi}{N\Delta k} \approx 1.57$ Å, where the size of the pixel in the reciprocal space, $\Delta k = 0.0074$ Å^{-1}, and the linear number of pixels, $N = 540$ pixels, were set by the experimental conditions. Unambiguous location of the geometrical centre of the orbital distribution in Fig. 6.2(A) is difficult.

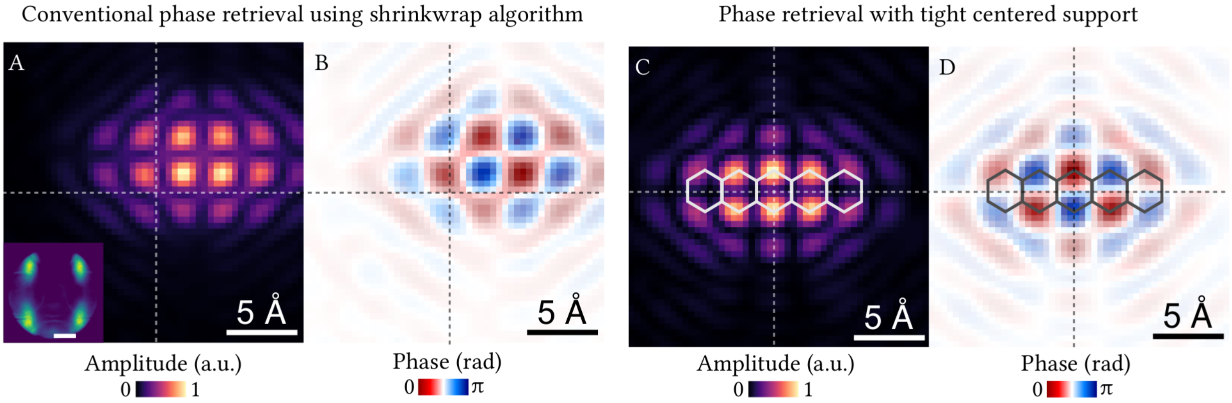


Figure 6.2: Reconstruction of the pentacene HOMO. (A) – Amplitude and (B) – Phase distributions reconstructed with the shrinkwrap algorithm using the uncentered support constraint. Inset in (A): The final PAD used as an input for the reconstruction procedure. The length of the scale bar corresponds to 1 Å^{-1}. (C) – Amplitude and (D) – Phase distributions reconstructed using the centred tight support obtained from the amplitude distribution in (A). The transparency of the phase images is weighted with the corresponding amplitude values for illustration purposes. The crossing lines mark the geometrical centres of the computational domains. Images (A-D) are 50×50 pixels sections cut out from 2000×2000 pixels reconstructed images. Adapted from [PK2].

6.3.2 *Phase retrieval with tight centred support*

To solve this problem, we undertook an additional reconstruction series, in which we employed the following support constraint:

1. The reconstructed amplitude distribution shown in Fig. 6.2(A) was centred in the computational domain using the central symmetry considerations.
2. The centred amplitude distribution was convoluted with a Gaussian function (standard deviation $\sigma_0 = 1.3$ pixels) and the resulting image was thresholded at 21% of its maximum. The threshold value was determined empirically so that the object distribution was not inadvertently truncated during the course of the reconstruction and the algorithm still converged.
3. The thresholded amplitude distribution was symmetrised with respect to its geometrical centre and the pixel values below 1 were set to 0, thus giving us a new object support.

The support was kept steady in the centre of the computational domain during the reconstruction. Similarly, we performed 1000 independent reconstruction rounds and each round consisted of 20 iterations of the PC-HIO algorithm followed by 20 iterations of the ER algorithm. Only 10% of the reconstructions with the lowest error metric were selected and averaged. The resulting amplitude and phase distributions of the pentacene HOMO are shown in Fig. 6.2(C,D). We note that the amplitude distribution in Fig. 6.2(A) shows two nodal planes in the centre of the support, while in Fig. 6.2(C) there is only one nodal plane. This is an example of a reconstruction artefact when the the correct reconstruction can be achieved only using the centred tight support.

6.4 DENSITY FUNCTIONAL THEORY SIMULATION

The results of the reconstruction we compared with the simulated orbital distributions. Electronic structure calculations of the pentacene HOMO in free space were performed using Kohn-Sham density functional theory (DFT) [16, 17] as implemented in the *Gaussian* quantum chemistry suite [153]. The three-dimensional (3D) orbital distribution was centred in the computational domain. The size of the pixel was $\Delta r = 0.1538$ Å in each dimension and the domain was sampled at $N_{x_0} \times N_{y_0} \times N_{z_0} = 51 \times 140 \times 72$ pixels. The 3D orbital distribution was zero-padded to $N = 512$ pixels in each dimension, yielding the orbital distribution $\psi(x, y, z)$. Fig. 6.3(A) shows an orbital isosurface plotted at 50% of the maximal value of $|\psi(x, y, z)|$ and three slices made at selected planes. The Fourier transform of $\psi(x, y, z)$ delivered the distribution $\Psi(k_x, k_y, k_z)$. Fig. 6.3(B) shows an isosurface plotted at 50% of the maximal value of $|\Psi(k_x, k_y, k_z)|$. The effective size of the pixel in the reciprocal space was equal to $\Delta k = \frac{2\pi}{N\Delta r} \approx 0.08$ Å^{-1}.

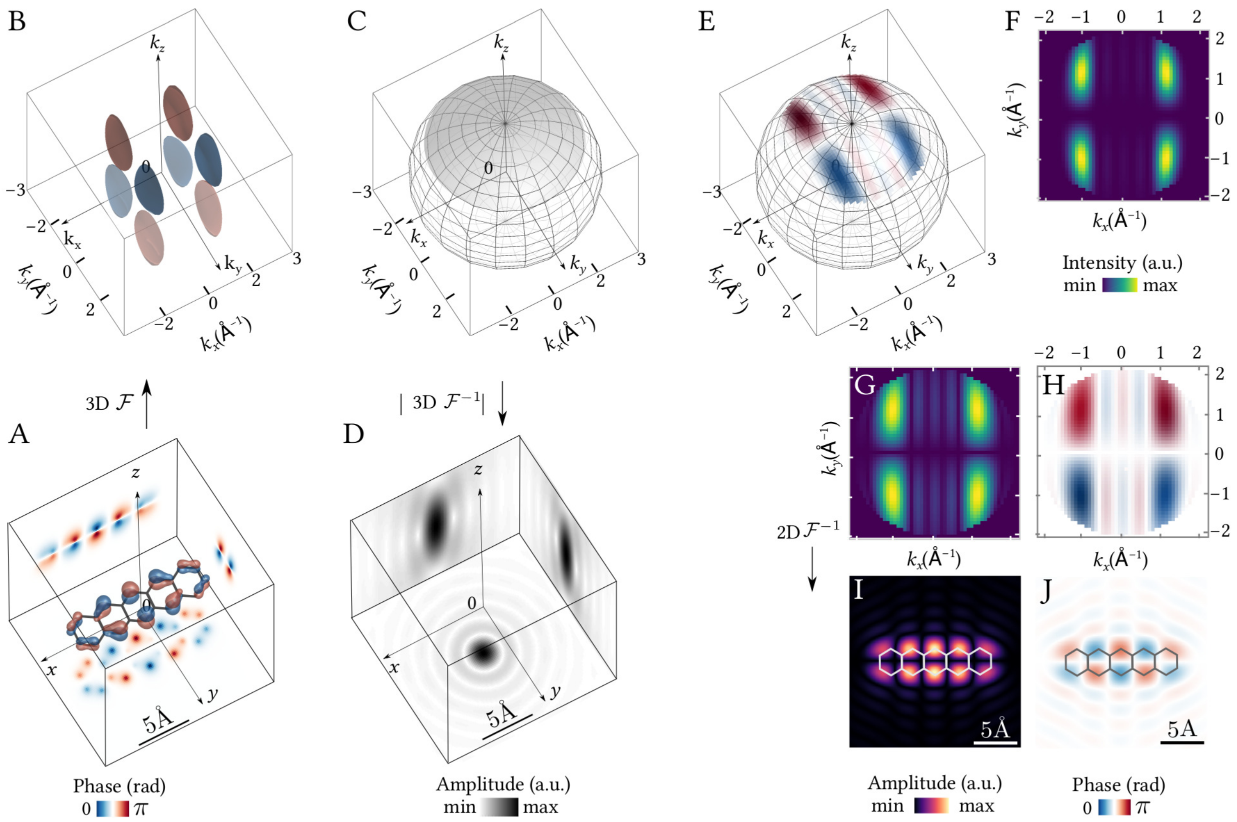


Figure 6.3: (A,B) – Results of the DFT simulation of the pentacene HOMO. (A) – $\psi(x,y,z)$: 3D orbital distribution, represented as isosurface at the value of 50% of maximum of $|\psi(x,y,z)|$. Inset images on the sides of the cube: Cross-sections through $\psi(x,y,z)$, computed at the planes located at $x \approx 2.61$ Å, $y \approx 1.54$ Å and $z \approx$ -0.15 Å. The phase values of the cross-sections are weighted with the corresponding amplitude values of $\psi(x,y,z)$ for illustration purposes. (B) – $\Psi(k_x,k_y,k_z)$: Fourier transform of $\psi(x,y,z)$, represented as an isosurface at the value of 50% of the maximum of $|\Psi(k_x,k_y,k_z)|$. (C-J) – Simulation of the experimental conditions using the results of the DFT simulation shown in (A,B). (C) – Transfer function $H(k_x,k_y,k_z)$: $H = 1$ on a segment of a hemisphere of radius $k_0 = 3.0$ Å^{-1} within the field of view of the parallel components of the momenta $k_{||,\max} = \pm\, 2$ Å^{-1}. $H = 0$ elsewhere. (D) – Inset images on the sides of the cube: Cross-sections through the amplitude of the response function $h(x,y,z) = \mathcal{F}^{-1}\{H(k_x,k_y,k_z)\}$, computed at the planes located at x = 0.0 Å, y = 0.0 Å and z = 0.0 Å. (E) – $\Psi(k_x,k_y,k_z)$ multiplied with the corresponding values of the transfer function H. (F) – Squared modulus of the parallel projection of (E) onto the (k_x,k_y) plane. To be compared with the experimental PAD shown in Fig. 6.1(F). (G, H) – Amplitude and phase distributions of the parallel projection of (E) onto the (k_x,k_y) plane. (I,J) – Amplitude and phase distributions in real space obtained by computing the inverse Fourier transform of (G) and (H). The phase values are weighted with the corresponding amplitude values for illustration purposes. To be compared with the reconstructed orbital distributions shown in Fig. 6.2(C) and (D). In (A,D), the grey borders of the cubes mark 110×110×110 pixels sections cut out from 512×512×512 pixels DFT data. In (B,C,E), the grey borders of the cubes mark 75×75×75 pixels sections cut out from 512×512×512 pixels DFT data. Adapted from [PK2].

To provide means for the quantitative comparison of the DFT results with the results of the iterative reconstruction and the experiment, we did a simulation by modifying the distribution $\Psi(k_x,k_y,k_z)$ so that it complied with the experimental conditions. The procedure was implemented in MATLAB.

1. In the experiment, the PAD was recorded at fixed kinetic energy. As optical transitions are direct transitions in reciprocal space, the momenta in the PAD were sampled on a hemisphere [1, 2, 4] with radius k_0 set by the photoelectron kinetic energy to $k_0 = 0.512 \cdot \sqrt{E_{\text{kin}}}$. At $E_{\text{kin}} = 34.3$ eV, the radius was $k_0 \approx 3.0$ Å^{-1}. Thus, all pixel values in the range of $[0.95k_0,\ 1.05k_0]$ were kept and all others were set to 0.

2. The values of all pixels lying outside of the $|k_{||}|_{\max} = 2$ Å^{-1} range were set to 0, in order to account for the numerical aperture of the electron entrance optics of the PEEM.

The choice of the $[0.95k_0,\ 1.05k_0]$ range was justified by two reasons. First, due to a limited number of pixels per reciprocal length unit, keeping the pixel values in this range ensured that no pixels were missing on the surface of the hemisphere. Second, this accounted for the (lorentzian) broadening of the photoelectron momentum perpendicular to the surface due to the exponential decay of the electron wave function inside the solid [154]. The decay length is essentially given by the inelastic mean-free path, which is of the order of 1 nm. This decay translates in a lorentzian momentum broadening of about 0.1 Å^{-1}, which agrees well with the interval chosen above.

The modification steps described above corresponded to multiplication of the $\Psi(k_x, k_y, k_z)$ distribution with a transfer function $H = H(k_x, k_y, k_z)$:

$$H = \begin{cases} 1 & \text{for } k_z = \sqrt{k_0^2 - k_x^2 - k_y^2} \cap k_x^2 + k_y^2 \leq k_{||,\max}^2 \\ 0 & \text{elsewhere.} \end{cases} \tag{6.11}$$

It represents a thin segment on the k-sphere as plotted in Fig. 6.3(C). Cross-sections through the amplitude of its complex-valued inverse Fourier transform done at selected planes are shown in Fig. 6.3(D). The resulting complex-valued distribution $\Psi_H\left(k_x, k_y, \sqrt{k_0^2 - k_x^2 - k_y^2}\right) = \Psi(k_x, k_y, k_z)H(k_x, k_y, k_z)$ is shown in Fig. 6.3(E). The squared modulus of this distribution, projected parallelly onto the (k_x, k_y) plane, shown in Fig. 6.3(F), corresponds to the intensity distribution measured in the experiment and is in good agreement with the PAD shown in Fig. 6.1(F). Amplitude and phase of $\Psi_H\left(k_x, k_y, \sqrt{k_0^2 - k_x^2 - k_y^2}\right)$ are shown in Fig. 6.3(G,H).

The inverse Fourier transform of the complex-valued distribution $\Psi_H\left(k_x, k_y, \sqrt{k_0^2 - k_x^2 - k_y^2}\right)$ yields a 2D orbital distribution in real space:

$$\psi'(x, y) = \iint_{-|k_{||}|_{\max}}^{+|k_{||}|_{\max}} \Psi_H\left(k_x, k_y, \sqrt{k_0^2 - k_x^2 - k_y^2}\right) e^{ik_x x + ik_y y} dk_x dk_y, \tag{6.12}$$

where the prime symbol distinguishes it from the original DFT data. Amplitude and phase of $\psi'(x, y)$ are shown in Fig. 6.3(I,J). These distributions mathematically correspond to those obtained by the iterative reconstruction shown in Fig. 6.2(C,D). We find them to be in good agreement, with some minor differences in the shapes of the individual lobes. The larger spatial extent of the reconstructed orbital distribution in Fig. 6.2 can be attributed to the side effects of the image processing procedure: in the processed experimental PAD, shown in Fig. 6.1(F), the blobs are more confined than those in the 2D distribution shown in Fig. 6.3(F) and the features in the interstitial area between the blobs disappear after the processing because they lie below the noise level. Consequently, the reconstructed orbital distribution becomes more delocalised in space.

6.4.1 *Interpretation of 2D orbital distributions and 3D reconstruction from 2D experimental data*

In Eq. 6.12, we employed the inverse Fourier transform to compute the 2D orbital distribution $\psi'(x, y)$ from the complex-valued photoelectron distribution $\Psi_H\left(k_x, k_y, \sqrt{k_0^2 - k_x^2 - k_y^2}\right)$. In essence, the same equation was employed while reconstructing the orbital distributions from the photoemission data. However, following the considerations of Barton [18], a more general expression for the orbital distribution ψ' is appropriate:

$$\psi'(x, y, z) = \iint_{-|k_{||}|_{\max}}^{+|k_{||}|_{\max}} \Psi_H\left(k_x, k_y, \sqrt{k_0^2 - k_x^2 - k_y^2}\right) e^{iz\sqrt{k_0^2 - k_x^2 - k_y^2}} e^{ik_x x + ik_y y} dk_x dk_y, \tag{6.13}$$

which we will refer to as "Barton's integral" [18]. The resulting reconstruction at $z = 0$ is shown in Fig. 6.3(I,J). If the integral in Eq. 6.13 is computed at some $z = z_0 = \text{const}$, then the resulting 2D orbital distribution $\psi'(x, y)$ is a superposition of the "in-focus" contribution $\psi(x, y, z_0)$ and the "out-of-focus" signal from adjacent z-planes as defined by the axial resolution of the experimental system.

The range of z planes contributing to the "out-of-focus" component can be estimated using the formula for the axial resolution defined by the Rayleigh range [37]:

$$\delta z = \frac{2\lambda_{\mathrm{dB}}}{\mathrm{NA}^2}. \tag{6.14}$$

Given the de Broglie wavelength of electrons at 34.3 eV kinetic energy is $\lambda_{\mathrm{dB}} \approx 2.1$ Å and the numerical aperture defined by the geometry in Fig. 6.3(C) is $\mathrm{NA} = \frac{|k_{||,\mathrm{max}}|}{k_0} = 0.67$, the axial resolution then equates to $\delta z \approx 9.4$ Å. The transverse resolution can be computed using the formula for the Airy radius [37]:

$$\delta r_{||} = \frac{0.61\lambda_{\mathrm{dB}}}{\mathrm{NA}}, \tag{6.15}$$

giving $\delta r_{||} \approx 1.9$ Å. These results are in a good agreement with the corresponding values estimated by computing the square of the amplitude of the response function $h(x, y, z)$ shown in Fig. 6.3(D).

Another important point we would like to highlight is the following. From the phase retrieval, we recover the complex-valued distribution $\Psi_H\left(k_x, k_y, \sqrt{k_0^2 - k_x^2 - k_y^2}\right)$. Now we note that in Eq. 6.13, the factor $e^{iz\sqrt{k_0^2-k_x^2-k_y^2}}$ plays the role of a propagator along the z-dimension of the orbital distribution. Thus, by computing Barton's integral at various values of z, one gains access to different "in-focus" contributions of the 3D orbital distribution $\psi(x, y, z)$ estimated at the planes z and one could thereby reconstruct the full 3D orbital distribution solely from the 2D distribution $\Psi_H\left(k_x, k_y, \sqrt{k_0^2 - k_x^2 - k_y^2}\right)$! This remarkable result was obtained by Barton in a simulated example of the holographic reconstruction of a S(1s) photoemitter signal from $c(2\times2)\mathrm{S/Ni}(001)$ data [18]. Though, for flat molecules with the thickness on the order of 2...3 Å, this will require improvement of axial resolution by increasing the photon energy and/or the numerical aperture of the system.

Using these results, we can also elaborate on reasons why orbital tomography was so far applied to planar molecules only [1, 2]. From one point of view, this limitation can be justified by the need of absence of scatterers on the way of the photoelectron wave as it propagates upon excitation from the molecule to the detector. However, another point of view can be gained from the interpretation of the 2D reconstruction as a superposition of "in-focus" and "out-of-focus" contributions. If a molecular orbital distribution is planar, the 2D reconstruction will contain "in-focus" and "out-of-focus" contributions having similar patterns in different z planes, blurred in accordance with the corresponding depth of field. The 2D orbital distribution thus represents a good estimate of the 3D orbital distribution. Otherwise, if the molecules are, for example, non-planar, the 2D reconstruction will contain "in-focus" and "out-of-focus" contributions of very different patterns from different z planes. The corresponding 2D reconstruction alone is thus no longer a good estimate of the 3D orbital distribution, but only some effective distribution defined by the orbital geometry and the depth of field. Unambiguous 3D reconstruction will thus require measuring PADs at multiple photon energies and solving the 3D phase problem, provided the oversampling requirements [9, 10] are fulfilled in all dimensions.

7

COBALT-PYRPHYRIN AND PYRPHYRIN ON AG(110)

The major part of this chapter will be submitted to a peer-review journal as “P. Kliuiev et al. Orbital tomography of complex macrocycles" [PK1].

In this chapter, we extend the orbital tomography technique to molecular systems relevant to catalysis: the bipyridine-based macrocycle pyrphyrin (Pyr) and its metalated form cobalt-pyrphyrin (CoPyr). The Pyr molecule was first synthesised by Ogawa et al. in 1984 [73] and is used as a carrier for catalytically active atoms of transition metals. CoPyr molecules are viewed as promising water reduction catalysts to produce hydrogen in photochemical water splitting [19, 20]. The ability to reconstruct molecular orbitals in such systems is important for the validation of theoretical models [20] and extension of orbital tomography to a broader class of molecules. In this work, we studied Pyr and CoPyr molecules adsorbed on Ag(110). ARPES measurements revealed the presence of several molecular states between the Fermi level and Ag d-band onset. For the purpose of orbital reconstruction with phase retrieval algorithms, the molecules must be well-ordered in a single rotational domain. Thus, it is important to know the spatial orientation of the molecules with respect to the substrate. The spatial mapping of molecular distributions can be done either by means of a scanning tunnelling microscopy (STM) or x-ray photoelectron diffraction (XPD) [155]. In this work, we employed an alternative method, based on a combination of LEED data, electronic structure calculations done using density functional theory DFT [16, 17] and orbital tomography. According to the results of the DFT simulation, CoPyr and Pyr molecules can adsorb on the surface of Ag(110) in at least three different conformations. The LEED results predict the existence of two molecular lattices for both types of molecular systems. Using this information and comparing the experimental PADs with those obtained using the DFT data within the plane wave (PW) approximation and the independent-atomic-centre (IAC) [6] approximation, we came to the conclusion that the CoPyr and Pyr molecules are ordered in three and two rotational domains on the surface of Ag(110), respectively.

7.1 CHARACTERISATION OF THE SAMPLES

The samples consisted of CoPyr and Pyr molecules adsorbed on Ag(110), prepared as described in chapter 4. For the characterisation of the system, we did three types of simulations:

1. DFT simulation of orbital distributions of CoPyr and Pyr molecules adsorbed on Ag(110)[1].

   Electronic structure calculations are performed at the Kohn–Sham DFT level using CP2K/ QUICKSTEP simulation package [156] in which Gaussian and plane wave formalism is implemented. While valence electrons are treated explicitly, the interaction between the valence electrons and the atomic cores is described through norm-conserving Goedecker–Teter–Hutter (GTH) pseudo potentials [157]. The number of valence electrons are set as 17, 4, 1, 5, and 11 for Co, C, H, N, and Ag, respectively. The molecular orbitals are expanded in Gaussian type orbitals using double-ζ plus polarisation (DZVP) basis sets, which have been optimised on molecular geometries (Mol-Opt method) [158]. The auxiliary plane wave basis used to represent the valence electron density in reciprocal space has an energy cutoff of 500 Ry. As exchange correlation functional, the spin polarised general gradient approximation by Perdew–Burke–Enzerhof (PBE) [159] is augmented with a nonlocal Vydrov–Voorhis van der Waals density functional, in its revised form [160] to account the dispersion contributions. The Ag bulk lattice parameter is optimised as 4.10 Å. The Ag surface is always described via a slab model, with the exposed plane perpendicular to the Cartesian axis. Periodic boundary conditions are always applied. To prevent any interaction between periodically replicated slabs, 20 Å of vacuum space are added along the surface normal. The study on the adsorption of the Pyr and CoPyr monomers is carried out using a 8x5 Ag(110) slab with five layers, where the four topmost layers are relaxed, whereas the fifth is kept fixed at the bulk coordinates.

---

[1] The DFT simulations of CoPyr and Pyr on Ag(110) were performed by Dr Y. Gurdal and PD Dr M. Iannuzzi, Department of Chemistry, University of Zurich.

To comply with the experimental conditions, the DFT data were modified as described in chapter 6 within the framework of the PW approximation.

2. DFT simulation[2] of orbital distributions of CoPyr and Pyr molecules in vacuum performed at the PBEPBE/cc-pVDZ level of theory [159, 161, 162] using the software "Gaussian" [153]. The DFT data were modified as described in chapter 6 within the framework of the PW approximation to comply with the experimental conditions.

3. Simulation of PADs for CoPyr and Pyr molecules in vacuum[3] within the framework of the IAC approximation using the results of DFT simulation done in 2. As an input, we used the output file of "Gaussian" [153]. The output file contained atom coordinates, the basis set and molecular orbital coefficients.

At first, the results of the DFT simulation of CoPyr and Pyr on Ag(110) were used together with the results of LEED to determine all possible spatial orientations of the molecules on the Ag(110) substrate. Then, this information was used in conjunction with the experimental data and the DFT and IAC simulations to establish which of the spatial orientations had contributed to the experimental PAD patterns.

DFT SIMULATION OF COPYR AND PYR ON AG(110) Fig. 7.1 shows the partial density of states (PDOS) and optimised adsorption geometries (AGs) with three lowest adsorption energies obtained from the DFT simulation of CoPyr and Pyr on Ag(110).

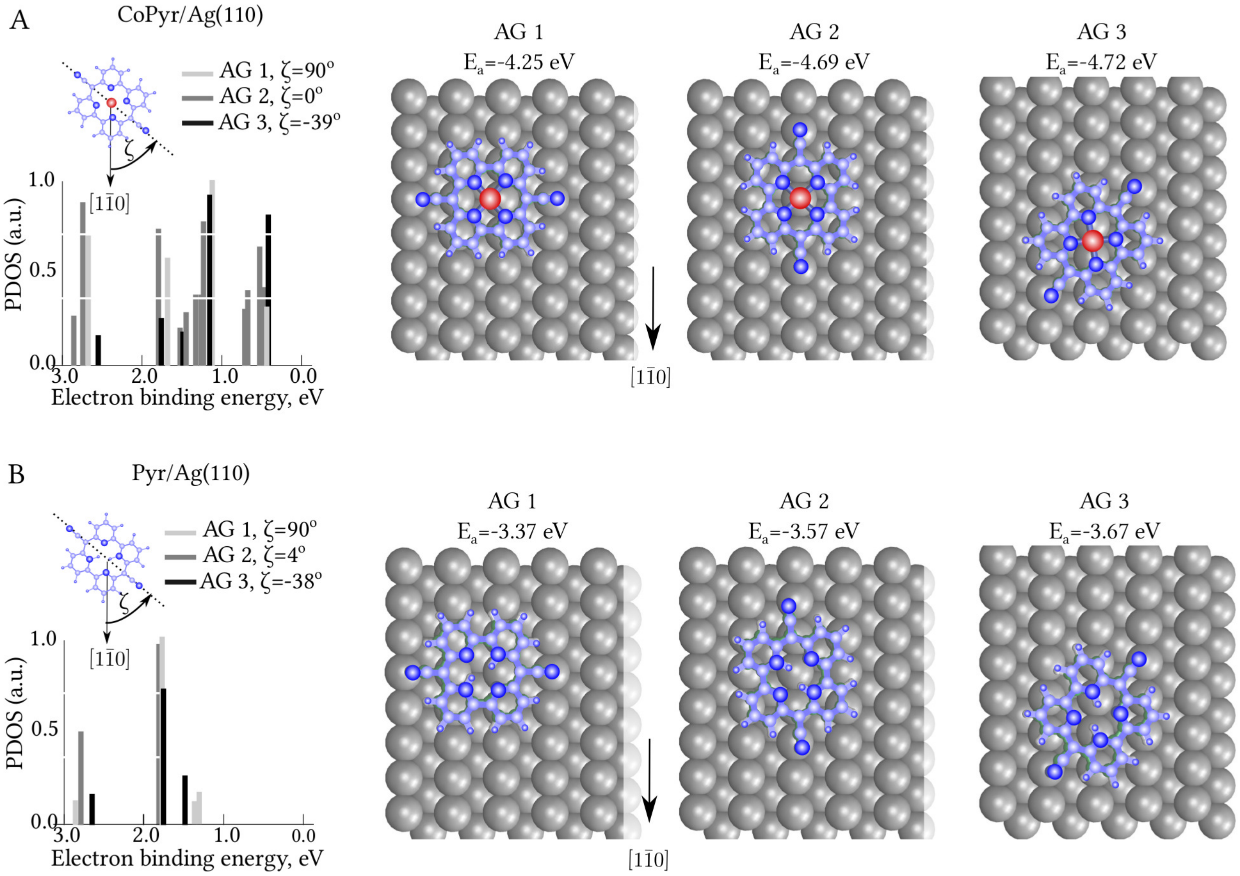


Figure 7.1: PDOS and optimised adsorption geometries of CoPyr and Pyr on Ag(110) obtained from DFT. Spatial orientation of molecules is specified by an angle ζ between the CN-groups and the $[1\bar{1}0]$ high symmetry direction of the substrate. Positive angles are measured counterclockwise. (A) – CoPyr on Ag(110). In the adsorption geometries (AGs) 1, 2 and 3, ζ are 90°, 0° and -39°, and the corresponding adsorption energies E_a are -4.25, -4.69 and -4.72 eV, respectively. (B) – Pyr on Ag(110). In the adsorption geometries (AGs) 1, 2 and 3, ζ are 90°, 4° and -38°, and the corresponding adsorption energies E_a are -3.37, -3.57 and -3.67 eV, respectively.

---

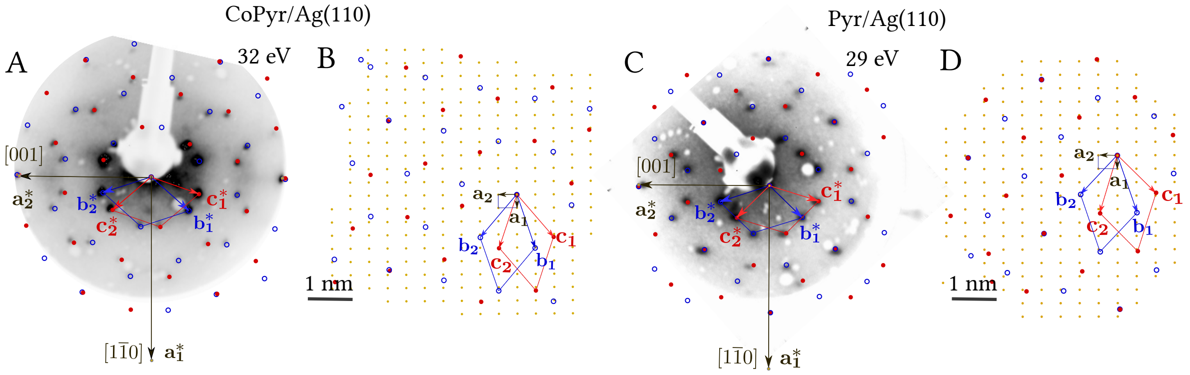
[2] The DFT simulation was performed by Dr L. Castiglioni, Department of Physics, University of Zurich.

[3] The simulations were done by the author of this dissertation using the software "IACPE-Calculation Tool" Version "18.0215 ScaSym BETA" developed by C. Metzger, University of Würzburg, Germany, implemented in the software IgorPRO.

It can be seen that the CoPyr molecules have slightly higher adsorption energies on the surface of the Ag(110) crystal compared to those of the Pyr molecules. This can be explained in terms of a more stable adsorption in the presence of a cobalt atom. For both CoPyr and Pyr on Ag(110), the most stable configuration with the lowest adsorption energy is AG 3, in which an angle between the CN-groups of CoPyr and Pyr and the $[1\bar{1}0]$ high symmetry direction of the substrate is $\zeta = -39°$ and $\zeta = -38°$, respectively. In the other two AGs 1 and 2, the angle between the CN-groups and the $[1\bar{1}0]$ high symmetry direction of the substrate is $\zeta = 90°$ and $\zeta = 0°$ for CoPyr and $\zeta = 90°$ and $\zeta = 4°$ for Pyr, respectively.

LEED OF COPYR AND PYR ON AG(110) To determine the epitaxial matrices of CoPyr and Pyr adlayers, the samples were characterised with LEED. Fig. 7.2(A,C) shows experimental LEED patterns of CoPyr and Pyr on Ag(110) together with the simulated patterns done in the framework of a geometric LEED theory using the commercial software LEEDLab (Scienta Omicron GmbH, Germany). Prior to the simulation, the experimental LEED images were corrected for barrel and Ewald sphere distortions using the procedure described in section 4. The simulation procedure consisted of the following steps:

1. The lattice constants of the Ag(110) substrate were set to $a_1 = 2.89$ Å and $a_2 = 4.09$ Å.
2. The software generated the lattices of molecular adlayers. The lattice parameters were set manually so that the simulated spot distributions approximately matched the experimental spots.
3. The experimental LEED spots were identified by a pattern recognition procedure of the software and fitted with Gaussian functions.
4. The mismatch between the lattice points generated in 2 and the experimental spots, fitted with Gaussian functions in 3, was minimised using the least squares method, in which the Ag(110) substrate lattice parameters defined in 1 were kept constant.
5. The optimised distribution of the fitted lattice structures are shown in Fig. 7.2(A,C). They were used to calculate the epitaxial matrices and the real space lattice structures shown in Fig. 7.2(B,D).



Reciprocal space Real space Reciprocal space Real space

Figure 7.2: (A,B) – CoPyr on Ag(110). (C,D) – Pyr on Ag(110). (A,C) – LEED patterns acquired at electron beam energies of 32 and 29 eV, respectively. Overlay: Simulated reciprocal lattice structures of the substrate (in grey) and of the two different lattices of molecular adlayers (lattice 1 and lattice 2, in blue and red, respectively). (B,D) – Real space lattice structures calculated from the optimised position of the simulated reciprocal lattice structures shown in (A,C). The colour scheme is identical to that in (A,C).

The LEED simulations yielded that CoPyr and Pyr molecules are ordered in two different lattices rotated with respect to each other. In Fig. 7.2(B,D), these lattices are shown in blue and red and we will refer to these lattices as "lattice 1" and "lattice 2", respectively. The corresponding epitaxial matrices and lattice parameters are given in Table 7.1. Within the error of measurement, the lattices parameters are identical and the lattices 1 and 2 are mirror counterparts of each other with respect to the $[1\bar{1}0]$ high symmetry direction.

	COPYR/AG(110)	PYR/AG(110)
Lattice 1 (blue)	$\begin{pmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{pmatrix} = \begin{pmatrix}4.15 \pm 0.08 & -0.97 \pm 0.05\\ 3.29 \pm 0.09 & 1.97 \pm 0.05\end{pmatrix}\begin{pmatrix}\mathbf{a}_1\\ \mathbf{a}_2\end{pmatrix}$ $\lvert\mathbf{b}_1\rvert = 12.59 \pm 0.23$ Å, $\lvert\mathbf{b}_2\rvert = 12.43 \pm 0.24$ Å, $(\widehat{\lvert\mathbf{b}_1\rvert, \lvert\mathbf{b}_2\rvert}) = -58.63^\circ \pm 0.04^\circ$	$\begin{pmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{pmatrix} = \begin{pmatrix}4.18 \pm 0.09 & -1.02 \pm 0.06\\ 2.96 \pm 0.09 & 1.96 \pm 0.06\end{pmatrix}\begin{pmatrix}\mathbf{a}_1\\ \mathbf{a}_2\end{pmatrix}$ $\lvert\mathbf{b}_1\rvert = 12.73 \pm 0.26$ Å, $\lvert\mathbf{b}_2\rvert = 11.69 \pm 0.25$ Å, $(\widehat{\lvert\mathbf{b}_1\rvert, \lvert\mathbf{b}_2\rvert}) = -62.24^\circ \pm 0.04^\circ$
Lattice 2 (red)	$\begin{pmatrix}\mathbf{c}_1\\ \mathbf{c}_2\end{pmatrix} = \begin{pmatrix}3.33 \pm 0.07 & -1.94 \pm 0.05\\ 4.14 \pm 0.07 & 1.00 \pm 0.04\end{pmatrix}\begin{pmatrix}\mathbf{a}_1\\ \mathbf{a}_2\end{pmatrix}$ $\lvert\mathbf{c}_1\rvert = 12.43 \pm 0.20$ Å, $\lvert\mathbf{c}_2\rvert = 12.60 \pm 0.20$ Å $(\widehat{\lvert\mathbf{c}_1\rvert, \lvert\mathbf{c}_2\rvert}) = -58.42^\circ \pm 0.03^\circ$	$\begin{pmatrix}\mathbf{c}_1\\ \mathbf{c}_2\end{pmatrix} = \begin{pmatrix}2.84 \pm 0.09 & -2.01 \pm 0.05\\ 4.25 \pm 0.10 & 0.95 \pm 0.06\end{pmatrix}\begin{pmatrix}\mathbf{a}_1\\ \mathbf{a}_2\end{pmatrix}$ $\lvert\mathbf{c}_1\rvert = 11.58 \pm 0.23$ Å, $\lvert\mathbf{c}_2\rvert = 12.84 \pm 0.28$ Å, $(\widehat{\lvert\mathbf{c}_1\rvert, \lvert\mathbf{c}_2\rvert}) = -62.65^\circ \pm 0.04^\circ$
Mutual orientation	$(\widehat{\lvert\mathbf{c}_1\rvert, \lvert\mathbf{b}_1\rvert}) = -21.21^\circ \pm 0.18^\circ$ $(\widehat{\lvert\mathbf{b}_1\rvert, [\mathbf{1\bar{1}0}]}) = 18.6^\circ \pm 0.1^\circ$ $(\widehat{\lvert\mathbf{c}_1\rvert, [\mathbf{1\bar{1}0}]}) = -18.6^\circ \pm 0.1^\circ$	$(\widehat{\lvert\mathbf{c}_1\rvert, \lvert\mathbf{b}_1\rvert}) = -26.01^\circ \pm 0.15^\circ$ $(\widehat{\lvert\mathbf{b}_1\rvert, [\mathbf{1\bar{1}0}]}) = 18.49^\circ \pm 0.1^\circ$ $(\widehat{\lvert\mathbf{c}_1\rvert, [\mathbf{1\bar{1}0}]}) = -17.00^\circ \pm 0.1^\circ$

Table 7.1: Epitaxial matrices and lattice parameters of real space lattices of CoPyr and Pyr on Ag(110). Substrate vectors are denoted by $\boldsymbol{a}_i$, those of the two different superstructure lattices by $\boldsymbol{b}_i$ and $\boldsymbol{c}_i$.

VISUALISATION OF MOLECULAR ROTATIONAL DOMAINS Using the information about the adsorption geometries and the surface lattices, it is now possible to visualise possible spatial orientations of the CoPyr and Pyr molecules on the Ag(110) surface. We set the distance between the CN-groups of CoPyr to 11.78 Å, as defined by the optimised molecular geometry in the gas phase [20]. Then, we note that because the lattice constants of the lattices 1 and 2 are on the order of the size of the molecule, it is not possible that lattices 1 and 2 co-exist simultaneously over a spatial extent on the order of a few lattice constants of the superstructure, as shown in Fig. 7.2(B,D). Hence, the two molecular lattices exist separately in the form of domains, defined by the three different adsorption geometries and the lattice structures 1 and 2, as shown in Fig. 7.3.

For symmetry reasons, both −39° and 39° adsorption geometries are possible. We will refer to them as "AG 3" and "AG 3′ ", respectively. Since the AG 3 leads to a very close separation between the CN-groups of the neighbouring molecules in the lattice 1 and the CN-groups tend to repel each other, this AG is excluded from consideration in lattice 1. The same argument applies to AG 3′ in lattice 2. Thus, in lattice 1, the CoPyr molecules can order themselves in the AGs 1, 2 and 3′ and in the lattice 2, in the AGs 1, 2, and 3, as shown in Fig. 7.3.

The separation between the CN-groups is the highest in AG 3 with the lowest adsorption energy and decreases with increasing adsorption energy, which correlates well with a notion that the CN-groups tend to repel each other and that the adsorption geometries with lower adsorption energies are more favourable.

Similarly, Fig. 7.4 shows possible spatial arrangements of the Pyr molecules. The distance between the CN-groups was set to 11.96 Å [20]. For symmetry reasons, both −38° and 38° as well as −4° and 4° adsorption geometries are possible. We will refer to them as "AG 3", "AG 3′ ", "AG 2′ " and "AG 2", respectively. Since the AG 3 leads to a very close separation between the CN-groups of the neighbouring molecules in the lattice 1 and the CN-groups tend to repel each other, this AG is excluded from consideration in lattice 1. The same argument applies to AG 3′ in lattice 2. The separation between the CN-groups in both lattices is large enough in the AGs 2 and 2′. Thus, in the lattice 1, the Pyr molecules can adsorb in the AGs 1, 2, 2′ and 3′, and in the lattice 2, in the AGs 1, 2, 2′ and 3, as shown in Fig. 7.4. Note that in Fig. 7.4(B,E), only the AG 2 is shown, while the AG 2′ is omitted. Similarly to CoPyr on Ag(110), the spacing between the CN-groups is the highest for AG 3 with the lowest adsorption energy and decreases with increasing adsorption energy, though, the rate, at which this happens, is slightly lower than in the case of CoPyr on Ag(110).

In essence, this analysis leads to the conclusion that, in principle, multiple rotational domains of the CoPyr and Pyr molecules on the surface of the Ag(110) substrate might potentially be present in the experiment. However, a final conclusion can be made only after the comparison of the experimental PADs with those obtained in the DFT and IAC simulations.

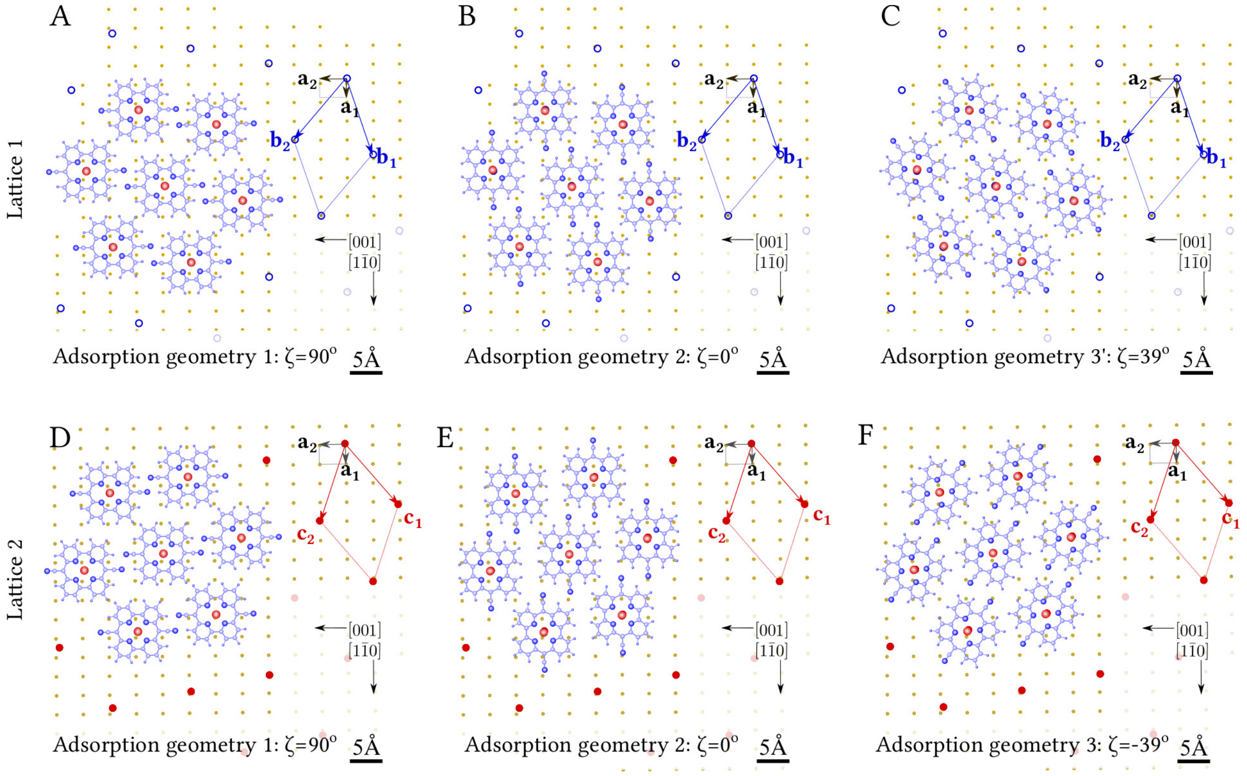


Figure 7.3: CoPyr molecules on Ag(110) ordered in multiple adsorption geometries in 2 lattice structures, as given by the DFT and LEED simulations, respectively. (A-C) – Lattice 1. (D-F) – Lattice 2. (A,D) – Adsorption geometry 1: $\zeta = 90°$. (B,E) – Adsorption geometry 2: $\zeta = 0°$. (C) – Adsorption geometry 3′: $\zeta = 39°$. (F) – Adsorption geometry 3: $\zeta = -39°$.

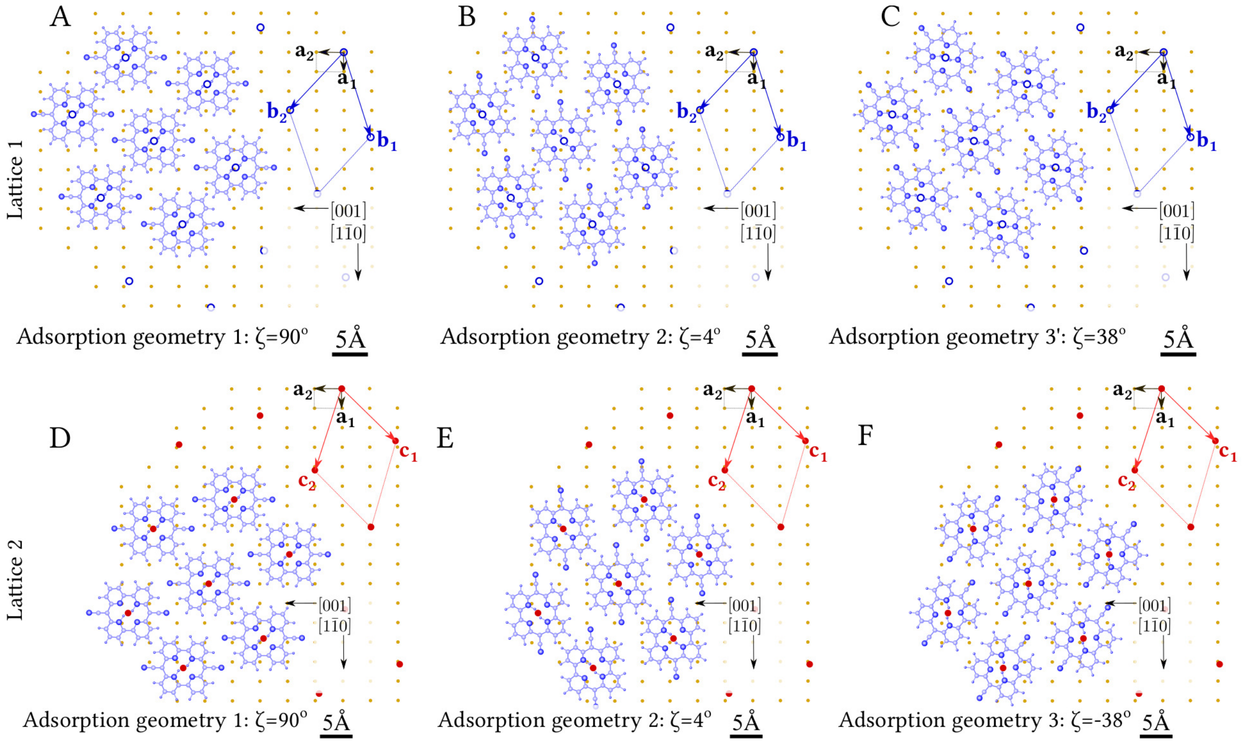


Figure 7.4: Pyr molecules on Ag(110) ordered in multiple adsorption geometries in 2 lattice structures, as given by the DFT and LEED simulations, respectively. (A-C) – Lattice 1. (D-F) – Lattice 2. (A,D) – Adsorption geometry 1: $\zeta = 90°$. (B,E) – Adsorption geometry 2: $\zeta = 4°$. In (B) and (E), adsorption geometry 2′ with $\zeta = -4°$ is also possible (not shown). (C) – Adsorption geometry 3′: $\zeta = 38°$. (F) – Adsorption geometry 3: $\zeta = -38°$.

7.2 ARPES

ARPES data of well-ordered monolayers of CoPyr and Pyr on the Ag(110) substrate were acquired with a photoemission electron microscope (PEEM) operating in momentum mode. The setup and the experimental geometry are shown in Fig. 4.2. The Ag(110) substrate was cleaned with standard sputter-anneal procedures and the molecules were evaporated in-situ using a home-built Knudsen cell evaporator as described in chapter 4.

Following the excitation of the samples with p-polarised light of 35 eV photon energy, we recorded photoelectron angular distributions (PADs) in steps of 50 meV at kinetic energies ranging from the onset of the Ag d-bands up to the Fermi level. By integrating the PADs over the whole field of view in momentum space at each of the recorded kinetic energies, we obtained the energy distribution curves (EDCs) shown in Fig. 7.5.

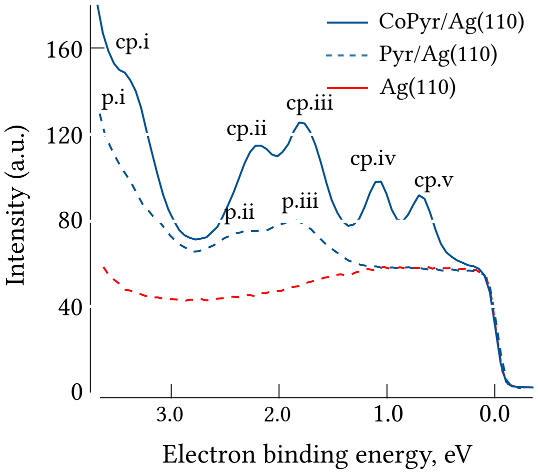


Figure 7.5: EDCs recorded from Pyr and CoPyr on Ag(110) and from the bare Ag(110) substrate using p-polarised light of 35 eV photon energy. Dwell time $t = 3$ s. The states of Pyr are marked as *"p.i"*, *"p.ii"* and *"p.iii"*, and those of CoPyr as *"cp.i"*, *"cp.ii"*, *"cp.iii"*, *"cp.iv"* and *"cp.v"*, from left to right.

Owing to the low density of states in Ag in this range of electron binding energies, the EDC recorded from the bare substrate did not reveal any substantial variations in the signal intensity. In contrast, the EDC recorded from Pyr on Ag(110) revealed 3 molecular states located at 3.50, 2.35 and 1.85 eV binding energies, which we will refer to as *"p.i"*, *"p.ii"* and *"p.iii"*, respectively. These can be seen as peaks in the EDC, though the peak at 3.50 eV is not clearly visible in the integrated data. In case of CoPyr on Ag(110), we could identify 5 states. Three of these states were located at 3.30, 2.20 and 1.80 eV binding energies, i.e. close to those found in the Pyr data, and the other two at 1.05 and 0.70 eV binding energies, respectively. We will refer to them as *"cp.i"*, *"cp.ii"*, *"cp.iii"*, *"cp.iv"* and *"cp.v"*, respectively. The PADs, recorded at kinetic energies of molecular states, are characterised by a number of blobs, i.e. intensity maxima, of different shape, size and mutual location, found between ±1...2 Å^{-1} as shown in Fig. 7.6(A,C). The corresponding PADs recorded from the bare substrate are shown in Fig. 7.6(B,D). Wide molecular features in Fig. 7.6(A,C) are superimposed with spurious crescent-like features originating from the sp-bands of the Ag(110) substrate.

7.3 DATA PROCESSING

To provide means for quantitative comparison with theory and for the iterative reconstruction with phase retrieval algorithms, the data shown in Fig. 7.6 (A,C), were processed: the background was subtracted and the distributions were normalised with the $|\boldsymbol{A} \cdot \boldsymbol{k}_f|^2$ factor. In chapter 6, we discussed means for the background subtraction using the maximisation of mutual information algorithm and a set of morphological operations. In this chapter, we subtracted the background in a more simple manner.

1. Experimental PADs acquired from CoPyr on Ag(110), shown in Fig.7.6(A,C), were centred in the computational domain, yielding the distributions $I_S(X,Y)$. The centre coordinates were found from the blurred PAD data, sampled at 540×540 pixels, upon convolution with a Gaussian function with the standard deviation of 3 pixels, using centrosymmetry considerations.

2. The signal distributions $I_S(X,Y)$ were registered with the corresponding Ag(110) background data, $I_B(X,Y)$, shown in Fig. 7.6(B,D), by manual application of affine transformations to the latter. For this purpose, the pixel values of $I_B(X,Y)$, sampled at 540×540 pixels, were weighted with a factor $s < 1$, the image was stretched by factors s_x and s_y in each dimension and the relative position of $I_B(X,Y)$ with respect to $I_S(X,Y)$ was adjusted by (t_x, t_y) pixels in horizontal and vertical directions, respectively:

$$I_b(X,Y) = s \cdot I_B(s_x X - t_x, s_y Y - t_y),$$

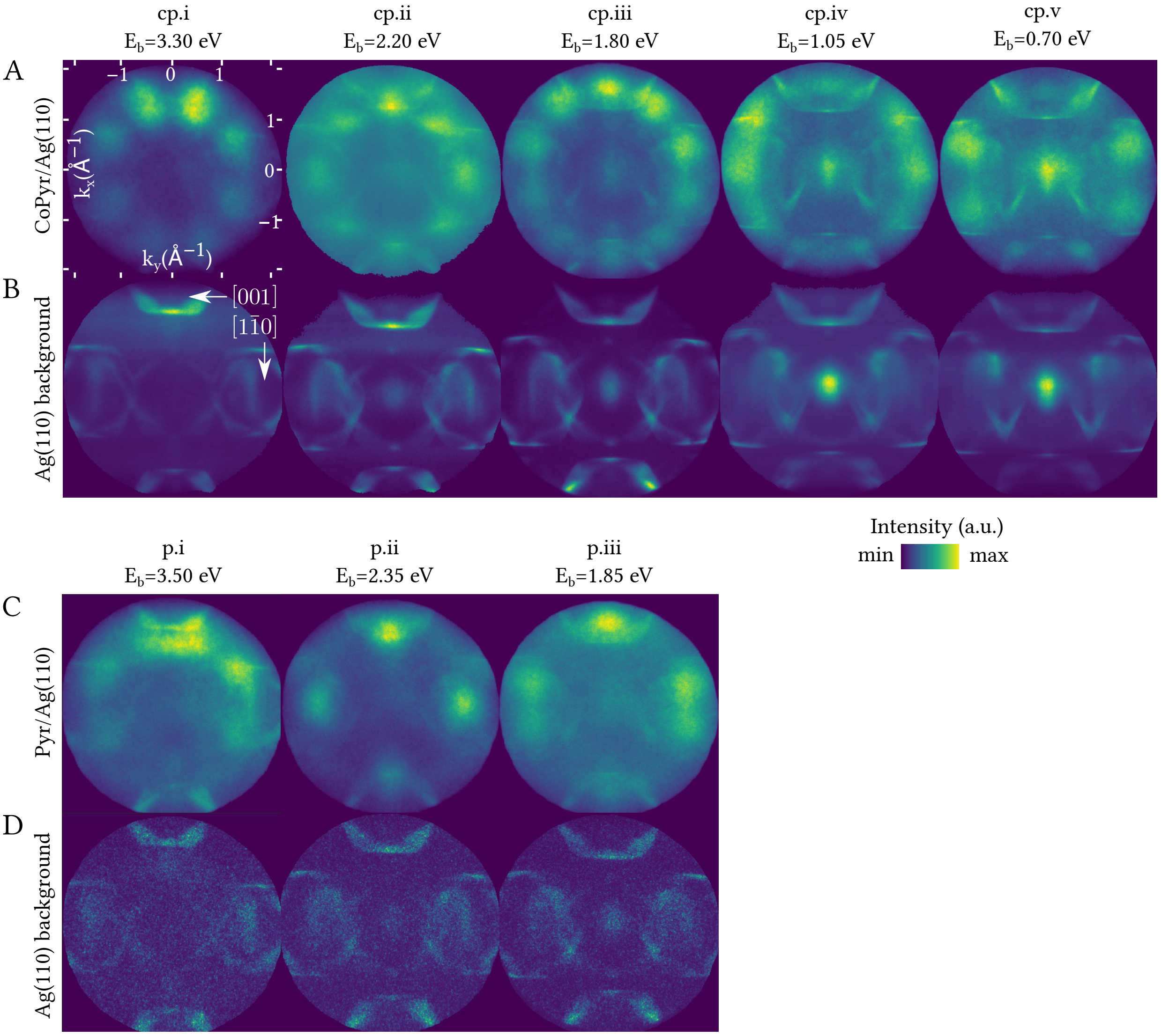
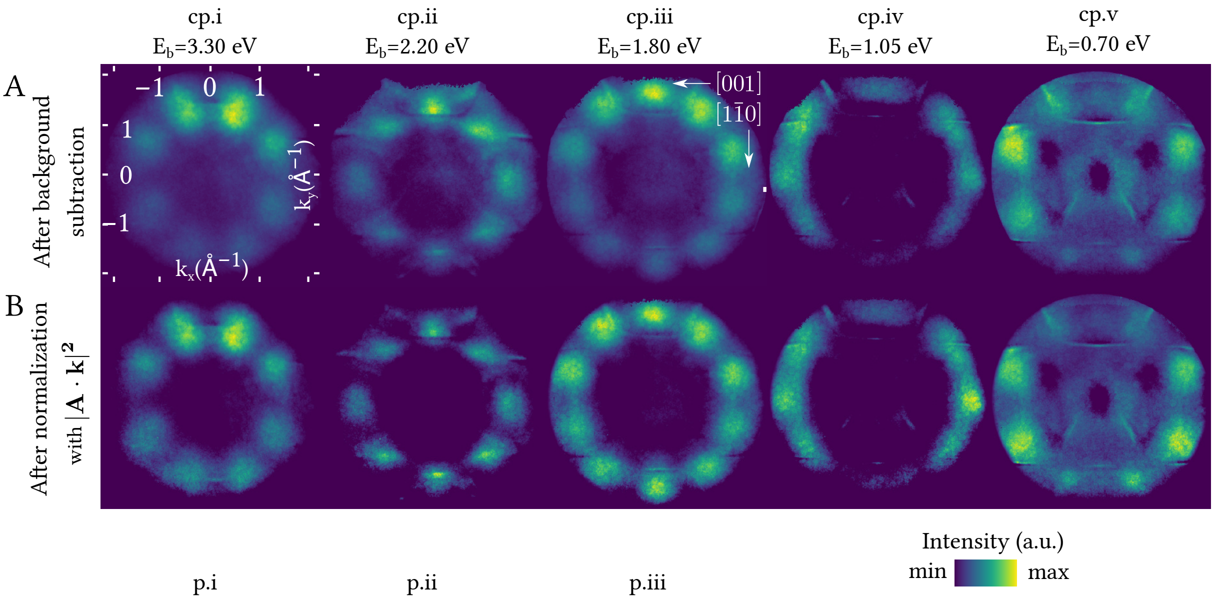


Figure 7.6: (A) – PADs recorded from CoPyr/Ag(110) with p-polarised light of 35 eV photon energy at the indicated binding energies of molecular states. Each momentum map is an average of 120 raw images acquired with a dwell time of 5 s each and aligned using the cross-correlation method described in chapter 6. (B) – PADs recorded from the bare Ag(110) substrate under the same conditions as in (A), except that each momentum map is an average of 50 raw images acquired with a dwell time of 3 s each. (C) – PADs recorded from Pyr/Ag(110) under the same conditions as in (A). (D) – PADs recorded from the bare Ag(110) substrate under the same conditions as in (A), except that each momentum map is a raw image acquired with a dwell time of 1 s.

where $I_b(X,Y)$ denotes the transformed background data. Initially, the parameters were set to $s = 1$, $s_x = s_y = 1$ and $t_x = t_y = 0$.

3. The background distribution $I_b(X,Y)$ was subtracted from $I_S(X,Y)$, yielding a new signal distribution: $I_s(X,Y) = I_S(X,Y) - I_b(X,Y)$.
4. The parameters in step 2 were adjusted and step 3 was repeated, until no further improvement in the quality of the distribution $I_s(X,Y)$ could be reached. The quality of the background subtraction was assessed by a visual inspection.



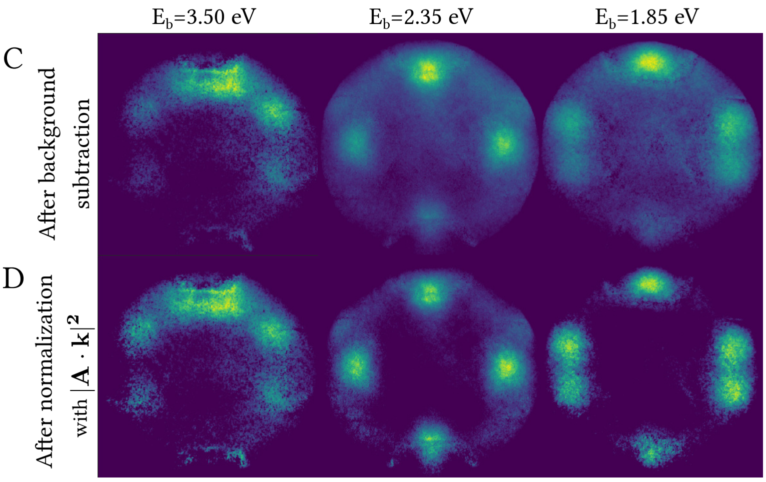


Figure 7.7: Processing of the experimental CoPyr/Ag(110) and Pyr/Ag(110) data. (A,B) – CoPyr/Ag(110). (C,D) – Pyr/Ag(110). (A,C) – PAD distributions after the background subtraction. (B,D) – After the normalisation with the $|\boldsymbol{A}\cdot\boldsymbol{k}_\mathrm{f}|^2$ factor.

The background-free PADs are shown in Fig. 7.7(A,C). Fig. 7.7(B,D) shows the PADs normalised by the $|\boldsymbol{A}\cdot\boldsymbol{k}_\mathrm{f}|^2$ factor, as described in chapter 6. For the purposes of comparison with the DFT data and reconstruction with phase retrieval algorithms, the data in Fig. 7.7(B,D) were symmetrised with respect to their geometrical centre and thresholded to eliminate the remnant background features. The symmetrised PADs are shown in Fig. 7.8(A.III) and 7.9(A.III), respectively.

7.4 COMPARISON OF THE EXPERIMENTAL PADS WITH THE DFT DATA

After subtraction of the background and normalisation, the experimental PADs were compared with those obtained from the DFT simulation within the PW and IAC approximations.

7.4.1 *CoPyr/Ag(110)*

DFT SIMULATION OF COPYR/AG(110). PW APPROXIMATION We started by comparing the experimental PADs with those obtained from the DFT simulation of CoPyr on Ag(110) within the framework of the PW approximation. For this purpose, the simulated 3D DFT orbital distributions were modified using the procedure described in chapter 6. All three adsorption geometries revealed the presence of very similar intensity distributions in reciprocal space that differed mostly only by their orientation with respect to the [1$\bar{1}$0] high symmetry direction of the substrate, as shown in Fig. 7.8(B.I-B.III). There was no agreement between the PADs in the AGs 1 and 3 and the normalised and symmetrised experimental data shown in Fig. 7.8(A.III), although AG 3 is the one with the lowest adsorption energy. The comparison of the PAD distributions in AG 2, which has got the second lowest adsorption energy, showed a better agreement. In particular, a very good agreement between the experimental and theoretical PADs was found for the molecular states *cp.i* and *cp.v*, while for the states *c.ii*, *c.iii* and *c.iv* the agreement was worse, as some of the blobs in the theoretical PAD distributions were found missing.

As it was indicated before, in principle, simultaneous existence of multiple adsorption geometries was possible. To verify this possibility, we performed a detailed analysis by combining the PADs in Fig. 7.8(B.I-B.III) and computing their incoherent sums. The results are shown in Fig. 7.8(B.IV-B.VI). Fig. 7.8(B.IV) shows the PADs obtained by computing a sum of the PADs in the AGs 3 and 3′. Among PADs in Fig. 7.8(B.IV), only those belonging to the molecular states *cp.i* and *cp.iv* agreed well with the corresponding experimental data. In contrast, much better agreement was found for the PADs obtained by performing a summation of the PADs in the AGs 2, 3 and 3′, weighted 1:1:1, as shown in Fig. 7.8(B.V). In this case, the number of blobs and their locations as well as mutual intensity distribution between the individual blobs matched those in the normalised and symmetrised experimental data. By varying the weighting coefficients, we found that the ratio 1.5:1:1 lead to the best agreement.

DFT SIMULATION OF COPYR IN VACUUM. PW APPROXIMATION The same procedure was applied to the DFT data for CoPyr in vacuum. We found the same molecular states as in the DFT data for CoPyr on Ag(110). We rotated the PADs by the corresponding angles ζ in azimuth so that they complied with the corresponding data in the adsorption geometries of CoPyr on Ag(110). We note that since a preferential orientation of molecules in vacuum does not exist, the sole purpose of this procedure is an azimuthal alignment of the data to provide means for the comparison with the data on the substrate. We will refer to the aligned data as "PADs in the adsorption geometries", although they are just PADs rotated by the corresponding angles ζ. PADs in the AGs 1 and 3 look similar to those in Fig. 7.8(B.I) and (B.III) and are not displayed here. Fig. 7.8(C.I) shows the PADs in the AG 2. Only the states *cp.i* and *cp.v* agreed well with the experimental PAD patterns in Fig. 7.8(A.III). As in the case of the DFT data for CoPyr on Ag(110), the patterns obtained by summation of the PADs in the AG 2, 3 and 3′, weighted 1.5:1:1, yielded the best agreement with the corresponding normalised and symmetrised experimental PADs in Fig. 7.8(A.III). We note that in the DFT data for CoPyr in vacuum, the mutual intensity distribution among the blobs was slightly different from that in the CoPyr/Ag(110) DFT data.

DFT SIMULATION OF COPYR IN VACUUM. IAC APPROXIMATION In the DFT simulation done within the framework of the IAC approximation, we set the adsorption geometries to be the same as for the data shown in Fig. 7.8(C.I-C.IV), i.e. the molecule was oriented at $\zeta = 0°$ and $\zeta = \pm 39°$ with respect to the $[1\bar{1}0]$ high symmetry direction of the substrate. We mimicked the experimental conditions, by setting the simulation parameters (photon energy, polarisation, angles of incidence) as in the experimental geometry shown in Fig. 4.2(B). The resulting PADs are shown in Fig. 7.8(D.I-D.IV). The PADs belonging to the states *cp.i* and *cp.v* in the AG 2 of the DFT data, looked very similar to those obtained after the background subtraction from raw experimental data shown in Fig. 7.8(A.II). In the PADs of the states *cp.ii*, *cp.iii* and *cp.iv*, some of the blobs present in the experimental data were found missing.

Analogously, the data obtained by summation of the PADs in the AGs 2, 3 and 3′, weighted 1.5:1:1, yielded the best agreement with the corresponding experimental data without background shown in Fig. 7.8(A.II). The agreement was particularly good for the molecular states *cp.i*, *cp.ii*, *cp.iii* and *cp.v*. The crescent-like features in the PADs of the states *cp.ii* an *cp.v*, were present both in the experimental and DFT-simulated data, and thus their origin was probably not only due to the substrate sp-bands. In the IAC approximation, the only disagreement was for the blobs in the upper part of the state *cp.iv*, which were slightly less prominent in the experimental data than in the DFT-simulated pattern.

The agreement between the experimental data and the DFT simulations within the framework of the PW and the IAC approximations, obtained by incoherent summation of the data in the adsorption geometries 2, 3 and 3′, is excellent. Neither of the single-domain adsorption geometries can reproduce all PAD features present in the experimental data. Thus, we conclude that CoPyr molecules on Ag(110) were adsorbed in three different geometries with the CN- groups oriented at 0°, 39° and −39° with respect to the $[1\bar{1}0]$ high symmetry direction of the substrate. This result correlates well with the analysis of the adsorption geometries done using the LEED data in Fig. 7.3 and with the fact that the AGs 3 and 3′ have the lowest adsorption energy $E_a = -4.72$ eV and the adsorption energy of AG 2 at $E_a = -4.69$ is only 30 meV higher.

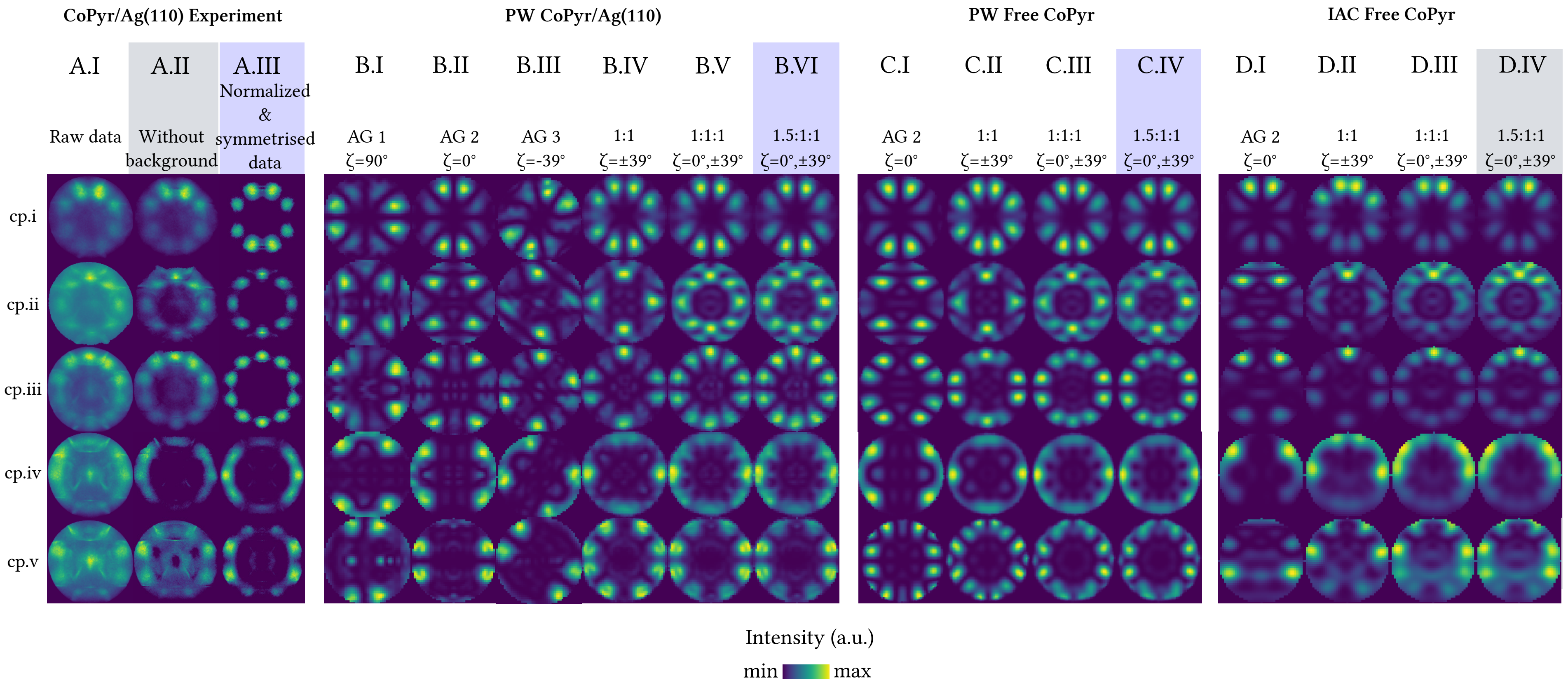


Figure 7.8: (A) – Experimental PADs recorded from CoPyr on Ag(110). (A.I) – Raw data. (A.II) – After background subtraction. (A.III) - After background subtraction, normalisation and symmetrisation. (B) – PADs obtained from the DFT simulation of CoPyr on Ag(110) within the PW approximation. (B.I) – Adsorption geometry 1: $\zeta = 90°$. (B.II) – Adsorption geometry 2: $\zeta = 0°$. (B.III) – Adsorption geometry 3: $\zeta = -39°$. (B.IV) – Sum of the data in the adsorption geometries 3 and 3′, weighted 1:1. (B.V) – Sum of the data in the adsorption geometries 2, 3 and 3′, weighted 1:1:1. (B.VI) – The same data as in (B.V), but weighted 1.5:1:1. (C) – PADs obtained from the DFT simulation of CoPyr in free space within the PW approximation. (C.I) – PADs with azimuthal orientation aligned to that in the adsorption geometry 2, $\zeta = 0°$, in (B.II). (C.II-C.IV) – Data obtained in the same way as in (B.IV-B.VI). (D) – PADs obtained from the DFT simulation of CoPyr in free space within the IAC approximation. (D.I) – PADs with azimuthal orientation aligned to that in the adsorption geometry 2, $\zeta = 0°$, in (B.II). (D.II-D.IV) – Data obtained in the same way as in (B.IV-B.VI).

7.4.2 *Pyr/Ag(110)*

We applied the same procedure to Pyr data.

DFT SIMULATION OF PYR/AG(110). PW APPROXIMATION We first compared the experimental PADs with those obtained from the DFT simulation of Pyr on Ag(110) within the framework of the PW approximation. As in the case of CoPyr on Ag(110), we found that all three adsorption geometries revealed the presence of very similar intensity distributions in reciprocal space that differed primarily only by their orientation with respect to the $[1\bar{1}0]$ high symmetry direction of the substrate, as shown in Fig. 7.9(B.I-B.III).

There was no agreement between the PADs in the AG 1 in Fig. 7.9(B.I) and the normalised and symmetrised experimental data in Fig. 7.9(A.III). The comparison of the experimental data in Fig. 7.9(A.III) with the AG 2 in Fig. 7.9(B.II) revealed an agreement only for the molecular state *p.i*, while for the molecular states *p.ii* and *p.iii* there was no agreement. Some of the PADs in AG 3 looked to us similar to those in Fig. 7.9(A.III) in terms of the number and location of the blobs. Though, a more careful comparison indicated that the best agreement was for an incoherent sum of the AGs 3 and 3′, weighted 1:1, shown in Fig. 7.9(B.IV). In this case, the number of blobs, their location and mutual intensities for all molecular states agreed well with those in the experimental data in Fig. 7.9(A.III).

DFT SIMULATION OF PYR IN VACUUM. PW APPROXIMATION The same procedure was applied to the DFT data for Pyr in vacuum. The DFT state *p.i* in Fig. 7.9(C.I) looked similar to the state *p.i* in the normalised and symmetrised experimental data shown in Fig. 7.9(A.III). Careful comparison showed that a better agreement is found for an incoherent sum of the AGs 3 and 3′, weighted 1:1, shown in Fig. 7.9(B.IV). The states *p.i*-*p.iii* agreed well in terms of the number of blobs, their location and mutual intensities with those in the experimental data shown in in Fig. 7.9(A.III).

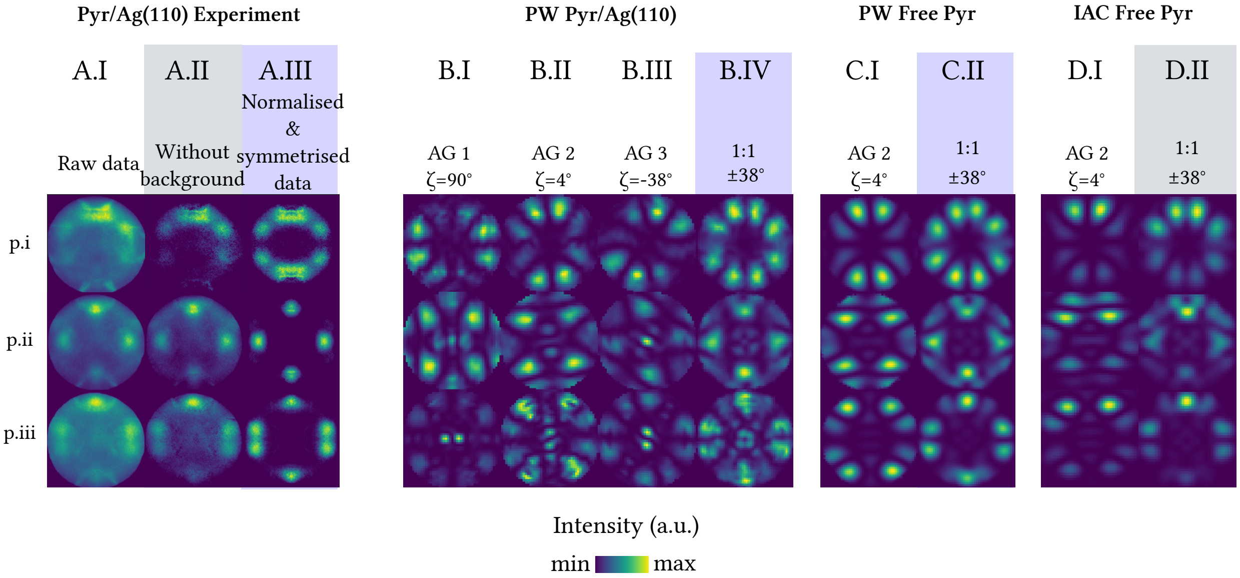


Figure 7.9: (A) – Experimental PADs recorded from Pyr on Ag(110). (A.I) – Raw data. (A.II) – After background subtraction. (A.III) – After background subtraction, normalisation and symmetrisation. (B) – PADs obtained from the DFT simulation of Pyr on Ag(110) within the PW approximation. (B.I) – Adsorption geometry 1: $\zeta = 90^\circ$. (B.II) – Adsorption geometry 2: $\zeta = 4^\circ$. (B.III) – Adsorption geometry 3: $\zeta = -38^\circ$. (B.IV) – Sum of the data in the adsorption geometries 3 and 3′, weighted 1:1. (C) – PADs obtained from the DFT simulation of Pyr in free space within the PW approximation. (C.I) – PADs with azimuthal orientation aligned to that in the AG 2 in (B.II). (C.II) – Sum of the PADs with azimuthal orientation aligned to that in the AGs 3 and 3′, weighted 1:1. (D) – PADs obtained from the DFT simulation of Pyr in free space within the IAC approximation. (D.I-D.II) – PADs modified in the same way as in (C.I-C.II).

IAC SIMULATION OF PYR IN VACUUM. IAC APPROXIMATION In the DFT simulation done within the framework of the IAC approximation, the data were treated in the same way as for CoPyr. The resulting PADs are shown in Fig. 7.9(D.I-D.II). Analogously, the data obtained by summation of the PADs in the

AGs 3 and 3′, weighted 1:1, revealed the best agreement with the corresponding experimental data in Fig. 7.9(A.II). The agreement was good for all molecular states in terms of number of blobs, their location and relative intensity distributions.

The agreement between the experimental data and the DFT simulations within the frameworks of the PW and the IAC approximations obtained by incoherent summation of the data in the adsorption geometries 3 and 3′ is excellent. Neither of the single-domain adsorption geometries can reproduce most of the PAD features present in the experimental data, as a sum of the AGs 3 and 3′. Thus, we conclude that Pyr molecules on Ag(110) were adsorbed in two different geometries with the CN- groups oriented at 38° and -38° with respect to the $[1\bar{1}0]$ high symmetry direction of the substrate. This result correlates well with the analysis of the adsorption geometries done using the LEED data in Fig. 7.4 and with the fact that the AGs 3 and 3′ have the lowest adsorption energy $E_a = -3.67$ eV.

7.5 RESULTS OF ITERATIVE RECONSTRUCTION

Despite of the fact that CoPyr and Pyr molecules were most possibly ordered on the surface of Ag(110) in three and two adsorption geometries, respectively, we performed an iterative reconstruction of distributions in real space by means of the phase retrieval algorithms. We employed the same reconstruction procedure as described in chapter 6. As an input, we used normalised and symmetrised PADs. The PADs from CoPyr and Pyr on Ag(110) shown in Fig. 7.8(A.III) and 7.8(A.III), used in phase retrieval, are duplicated in Fig. 7.10(A) and 7.12(A), respectively.

7.5.1 *CoPyr/Ag(110)*

At first, the reconstruction was done using the shrinkwrap algorithm [12]. The resulting real space distributions are shown in Fig. 7.10(B,C).

While it was simple to determine the geometrical centres of the states *cp.i-cp.iii*, it was not the case for the states *cp.iv-cp.v*. As it was discussed in chapter 6, these are examples of reconstruction artefacts and the correct reconstructions can only be achieved using a centred tight support. Fig. 7.10(D,E) shows the results of the reconstruction with the centred tight support obtained using the amplitude distributions in Fig. 7.10(B) together with the overlaid ball-and-stick models of the CoPyr molecule. Due to the different image processing procedure, the blobs in the PADs shown in Fig. 7.10(A) became more localised and, as a result, the reconstructed real space distributions became more delocalised. Also, we note that the overlaid ball-and-stick models are only guides for the eye and only mark the regions of space where the orbital distributions might potentially reside. The distributions shown in Fig. 7.10(B-E) represent only some real space distributions reconstructed from an incoherent sum of the PADs and their meaning will be discussed later in the text.

In spite of this ambiguity, we compared the results of the reconstruction with the DFT data in the AG 2 ($\zeta = 0°$), because they resembled the reconstructed data in terms of azimuthal orientation. We computed the corresponding real space orbital distributions using the procedure described in chapter 6. The resulting amplitude and phase distributions are shown in Fig.7.11(B,C,E,F).

Surprisingly, we found a good agreement for three molecular states - *cp.i*, *cp.iii* and *cp.iv* - in terms of the number of lobes, their location and mutual amplitude distribution among the lobes. The agreement for the states *cp.ii* and *c.v* was less prominent, owing to the differences in the amplitude distribution among different orbital lobes. For all molecular states, the relative phase distributions between individual lobes in the DFT simulation coincided with those in the reconstructed data.

7.5.2 *Pyr/Ag(110)*

We repeated the same procedure for the Pyr data. Fig. 7.12(A) shows the resulting PAD distributions after the background subtraction, normalisation and symmetrisation. Since the pixel values in the lower part of the state *p.i* were on the order of noise level and thus, in essence, became zero after the background subtraction, we mirrored the upper part of the PAD to obtain at least some signal distribution in the lower part of the PAD. For the iterative reconstruction, we employed the same procedure as described in chapter 6. The real space distributions reconstructed with the shrinkwrap algorithm [12] are shown in Fig. 7.12(B,C). It was not possible to obtain a meaningful reconstruction for the *p.i* data. Varying the algorithm parameters did not lead to the improvement of the reconstruction results. In the second reconstruction run, we employed a centred tight support obtained from the amplitude distributions in Fig. 7.12(B) as described in chapter 6. The use of the tight centred support improved the quality of the reconstructed orbital distributions. In particular, for the state *p.i*. The resulting real space distributions are shown in Fig. 7.12(D,E).

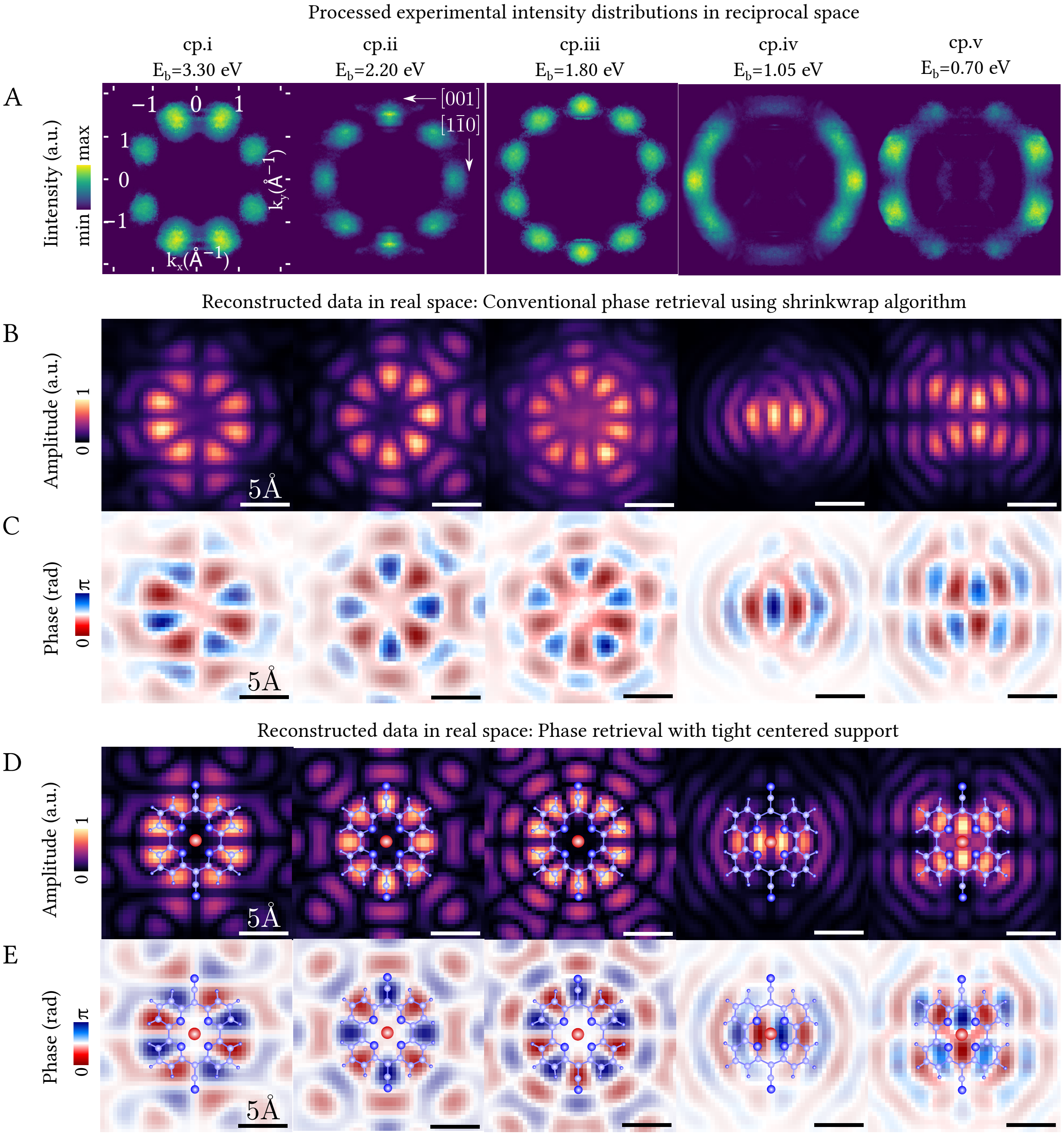


Figure 7.10: Results of iterative phase retrieval procedure applied to CoPyr/Ag(110) data. (A) – Experimental PADs after the background subtraction, normalisation and symmetrisation. (B,C) – Real space distributions reconstructed with conventional phase retrieval using shrinkwrap algorithm. (D,E) – Real space distributions reconstructed with the centred tight support. (B,D) – Amplitude distributions. (C,E) – Phase distributions weighted with the corresponding amplitude values for illustration purposes.

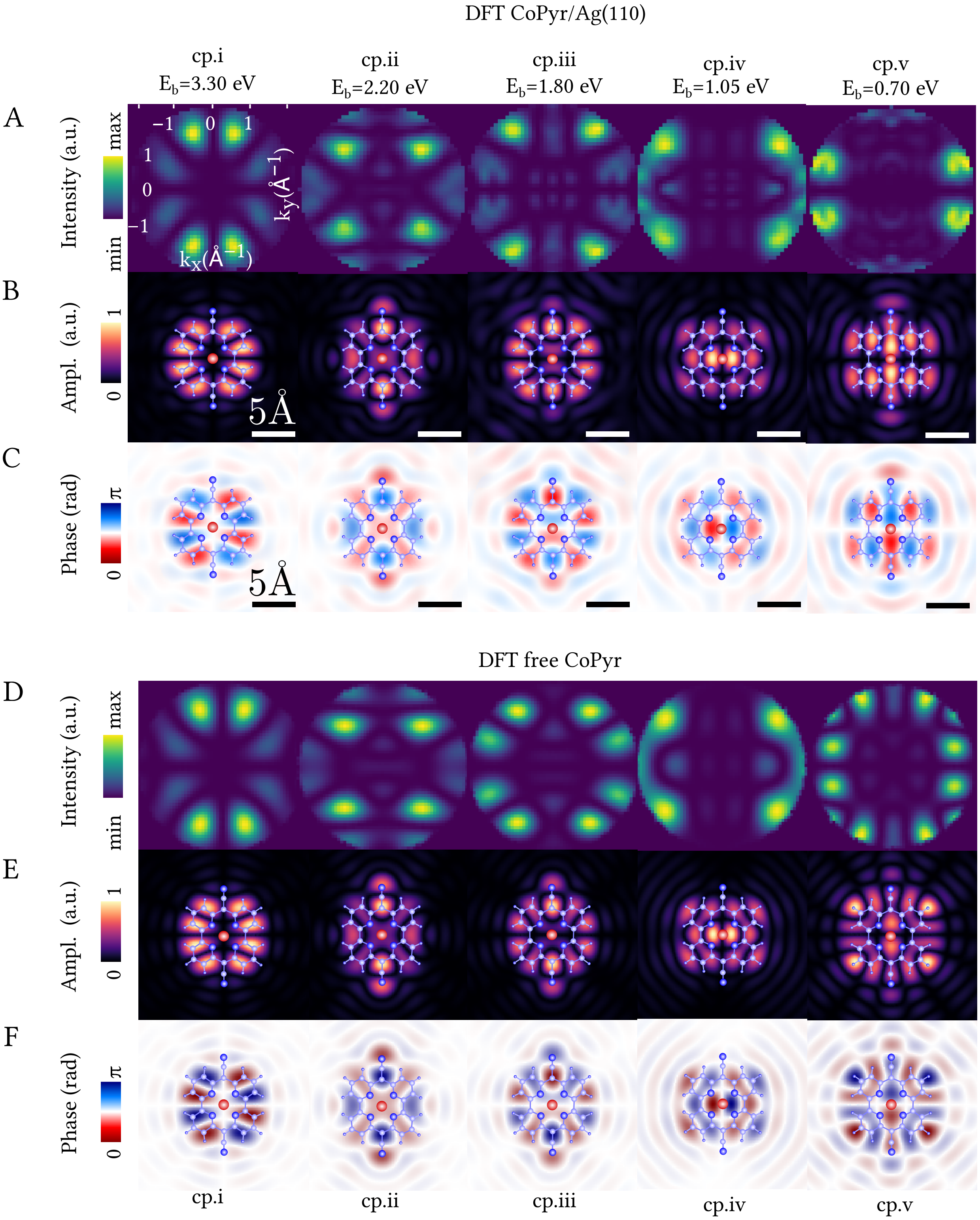


Figure 7.11: Results of the DFT simulation. (A-C) – CoPyr on Ag(110). Adsorption geometry 2: $\zeta = 0°$. (D-F) – CoPyr in free space. (A,D) – Intensity distributions in reciprocal space. (B,E) – Amplitude distributions in real space. (C,F) – Phase distributions in real space weighted with the corresponding amplitude values for illustration purposes. Orbital distributions are ordered by the respective binding energies of molecular states in the experiment.

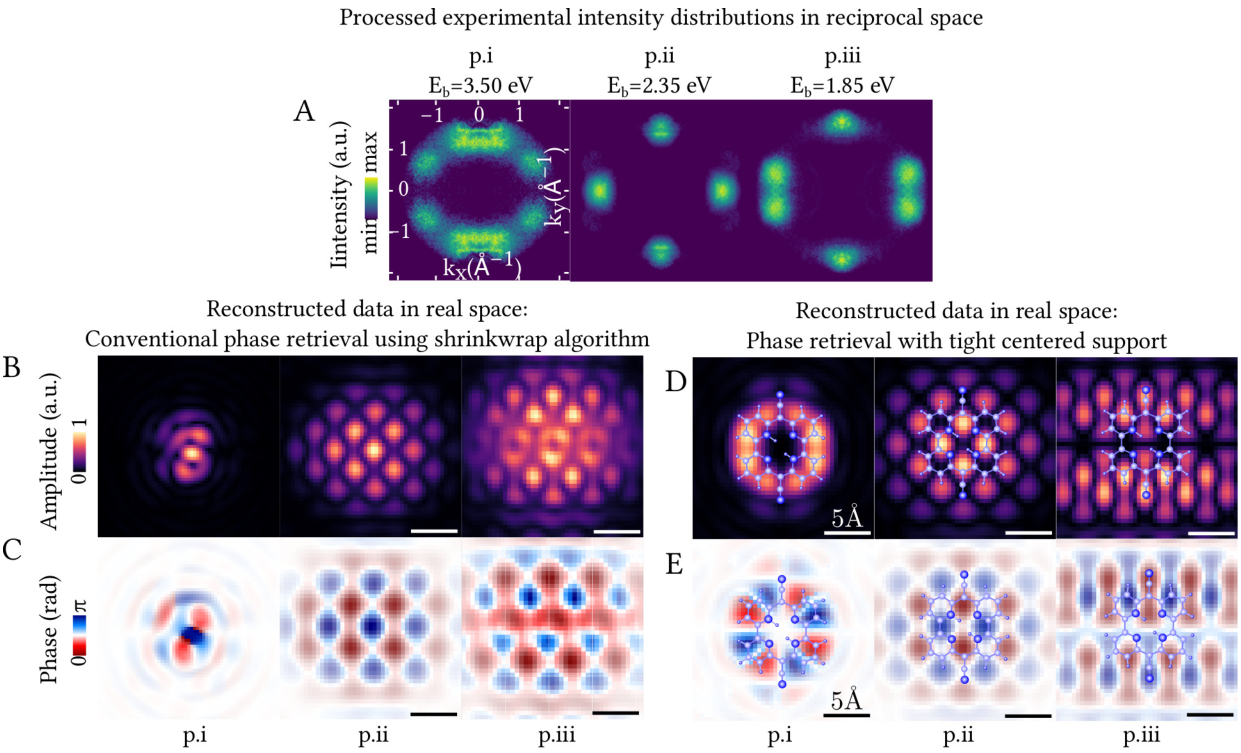


Figure 7.12: Results of iterative phase retrieval procedure applied to Pyr/Ag(110) data. (A) – Experimental PADs after background subtraction, normalisation and symmetrisation. (B,C) – Real space distributions reconstructed with conventional phase retrieval using shrinkwrap algorithm. (D,E) – Real space distributions reconstructed with the centred tight support. (B,D) – Amplitude distributions. (C,E) – Phase distributions weighted with the corresponding amplitude values for illustration purposes.

In Fig. 7.12(D,E), the ball-and-stick molecular models overlaid over the reconstructed real space distributions serve as guides for the eye only and mark the regions of space where the orbital distributions might potentially reside. The distributions in Fig. 7.12(D,E) represent only some real space distributions reconstructed from an incoherent sum of the PADs and their meaning will be discussed later in the text.

We compared the results of the reconstruction with the DFT data in the AG 2 ($\zeta = 4°$), because they resembled the reconstructed data in terms of azimuthal orientation. The real space distributions were computed using the procedure described in chapter 6. In the DFT data for Pyr in vacuum and on Ag(110), only the state *p.i* agreed well with the reconstructed data.

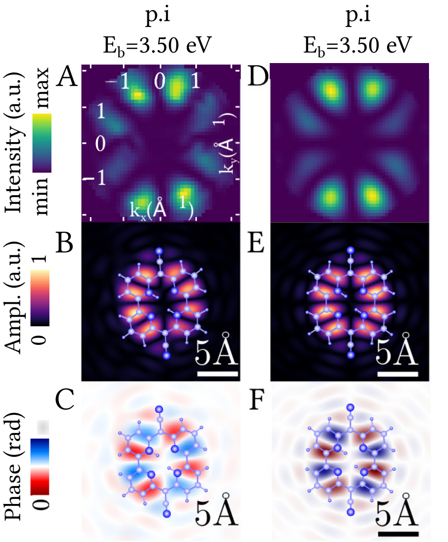


Figure 7.13: Results of the DFT simulation. (A-C) – Pyr on Ag(110) in free space. (D-F) – Pyr on Ag(110) in AG 2. (A,D) – Intensity distributions in reciprocal space. (B,E) – Orbital amplitude distributions. (C,F) – Orbital phase distributions weighted with the corresponding amplitude values for illustration purposes.

7.6 SUMMARY AND DISCUSSION

A detailed analysis of the LEED and DFT data in Fig.7.3 and 7.4 yielded that CoPyr and Pyr molecules are ordered on the surface of Ag(110) in two lattices and be adsorbed in multiple geometries. In lattice 1, CoPyr molecules can order themselves in the adsorption geometries 1, 2 and 3′, so that the CN- groups are located at ζ =90°, 0° and 39° with respect to the $[1\bar{1}0]$ high symmetry direction of the substrate, correspondingly. In lattice 2, CoPyr molecules can order themselves in the adsorption geometries 1, 2 and 3, so that the CN- groups are located at ζ =90°, 0° and −39° with respect to the $[1\bar{1}0]$ high symmetry direction of the substrate, correspondingly. Thus, CoPyr molecules on Ag(110) can be in principle ordered in 4 different rotational domains. Because the adsorption geometries 3 and 3′ have the lowest adsorption energy $E_a = -4.72$ eV, and the adsorption energy of the adsorption geometry 2 is just 0.03 eV higher, CoPyr molecules can adsorb in the adsorption geometries 2, 3 and 3′ with almost equal probability. Similarly, we found that Pyr molecules can order themselves in lattice 1 in the adsorption geometries 1, 2, 2′ and 3′, so that the CN- groups are located at ζ =90°, ±4° and 38° with respect to the $[1\bar{1}0]$ high symmetry direction of the substrate, correspondingly. In lattice 2, Pyr molecules can order themselves in the adsorption geometries 1, 2, 2′ and 3, so that the CN- groups are located at ζ =90°, ±4° and −38° with respect to the $[1\bar{1}0]$ high symmetry direction of the substrate, correspondingly. Thus, Pyr molecules on Ag(110) can be in principle ordered in 5 rotational domains. The adsorption geometries 3 and 3′ have the lowest adsorption energy $E_a = -3.67$ eV, while the adsorption energy of the geometries 2 and 2′ is 0.1 eV higher. Thus, Pyr molecules rather adsorb in the adsorption geometries 3 and 3′ only.

The results of the incoherent summation of the DFT-simulated PADs of the CoPyr and Pyr molecules on the Ag(110) substrate and in vacuum performed within the frameworks of the PW and IAC approximations confirmed our assumptions. CoPyr molecules revealed three adsorption domains, with the CN- groups located at $\zeta = 0°$ and ±39° with respect to the $[1\bar{1}0]$ high symmetry direction of the substrate. Pyr molecules ordered themselves in two domains, in which the CN- groups are located at ±38° with respect to the $[1\bar{1}0]$ high symmetry direction of the substrate.

Having obtained this result, we nevertheless performed reconstructions of real space distributions from the experimental data. Interestingly, we found the real space reconstructions to be in good agreement with the corresponding DFT orbital distributions in the AG 2, obtained from the simulations on the substrate and in vacuum. This agreement can be explained as follows. The experimental patterns recorded from CoPyr on Ag(110) are incoherent sums of the PADs from the adsorption geometries 2, 3 and 3′. Thus, they contain features from the PADs belonging to the AG 2, but also those belonging to the AGs 3 and 3′. We make a comparison with the DFT data in the AG 2. Obviously, the presence of the common features in reciprocal space (those which belong to the AG 2) makes the reconstructed and the DFT orbital distributions look similar. However, a very close comparison of the data reveals that, these distributions are not identical. For instance, the number of lobes and their relative strength in the reconstructed CoPyr data *cp.ii* in Fig. 7.10(D,E) is not exactly the same as in the DFT orbital distributions in Fig. 7.11(B,C,E,F). Additional lobes in real space have arisen due to the blobs present in the experimental data at around ± 1.5 Å^{-1}. These features are absent in the DFT data in reciprocal space and thus, in real space, the DFT-simulated orbital distributions look slightly differently from the reconstructed real space distributions. Similar reasoning can be made about other CoPyr states. A better insight into the similarity of the experimental PADs and the PADs simulated for a single rotational domain of the molecules, i.e. AG 2, can be gained by inspecting the symmetry properties of the latter. Because some of the blobs in the PADs in the AG 2 are separated by $\approx 40°$, rotation of PADs by ±39° and their summation with the AG 2, amplifies the signal at the positions of the blobs in the latter (e.g. compare the PADs of the *cp.i* state in Fig. 7.8(B.II - B.V)).

To summarise, the remarkable agreement between the reconstructed real space distributions and the DFT orbital distributions can be attributed to the presence of common features in the corresponding data in reciprocal space. And it is this favourable combination of the symmetry properties of individual PADs and molecular adsorption geometries that makes incoherent sum of the DFT-simulated PADs and experimental data look similar.

We note that for the phase retrieval algorithm, it plays no role if the blobs in the experimental PADs are due to photoemission from a single or multiple rotational domains of the molecules. Solving the phase problem, one will always be able to reconstruct some amplitude and phase distributions in real space, provided the oversampling conditions are fulfilled [9]. Currently, we cannot give any meaningful physical interpretation for real space distributions reconstructed from an incoherent superposition of intensity distributions in reciprocal space. To the best of our knowledge, unambiguous interpretation of reconstructed orbital distributions is possible only if the photoemission data originate from a single domain of molecules. Thus, the origin of the data used as an input for phase retrieval algorithms must always be clarified prior to any interpretation of reconstructed real space distributions.

Under certain conditions, perhaps, it might be possible to disentangle the PADs from single rotational domains directly from incoherent sums of the PADs from multiple rotational domains by means of independent component analysis (ICA) [163]. ICA is a statistical method that decomposes a complex dataset into its independent constituent parts. These parts have to be statistically independent, i.e. one of the pixel values of any one of the components must contain no information on the pixel values of the other components. In practice, this means that the experimental data must be recorded from multiple samples, in which molecular rotational domains have different coverages (for example, in one sample, 20% of molecules in one domain, 80% in the other and in another sample 50% of molecules in one domain, 50% in the other) and are non-interacting.

In an ideal case, of course, one should avoid multiple rotational domains of molecules, if the main goal is reconstruction of real space orbital distributions. The reconstruction process by means of phase retrieval requires unbiased input with minimal changes in the experimental data. Thus, even in case of a single rotational domain, the presence of the substrate background features makes the reconstruction process very complicated and biased, as it requires substantial amount of image processing. To conclude, for successful reconstruction of orbital distributions with phase retrieval, one should have ARPES data without substantial amount of background features, acquired with good statistics from a single rotational domain of the molecules. This requirement is particularly important, especially in case if one aims at time-resolved reconstruction of orbitals, because in this case we add an additional degree of freedom and orbital reconstruction would be greatly facilitated if the experimental data were unambiguously recorded from a well-defined molecular system.

Finally, we also would like discuss the role of Co atom in influencing the molecular orbitals of Pyrphyrin on Ag(110). Because CoPyr molecules order themselves on the surface of the Ag(110) crystal in three different rotational domains and Pyr molecules only in two, direct comparison of the experimental PADs is not possible. However, individual adsorption geometries simulated with the DFT reveal the presence of very similar intensity distributions in reciprocal space that differ only by their orientation with respect to the $[1\bar{1}0]$ high symmetry direction of the substrate, as shown in Fig. 7.8(B.II-B.III). Thus, it makes sense to compare the DFT-simulated PADs stemming from similar adsorption geometries of CoPyr and Pyr on Ag(110), i.e. from AG 3 of CoPyr and AG 3 of Pyr, in which the CN groups are oriented at $-39°$ and $-38°$ with respect to the $[1\bar{1}0]$ high symmetry direction of the substrate, respectively. From Fig. 7.8(B.III) and Fig. 7.9(B.III), it is seen that the PADs belonging to molecular states of Pyr, $p.i - p.iii$, and of CoPyr, $cp.i - cp.iii$, have got identical features, though some of them are more pronounced in case of CoPyr on Ag(110). The energy positions of CoPyr molecular states, $cp.i - cp.iii$, are slightly shifted to lower binding energies compared to the energy positions of the corresponding states of Pyr, $p.i - p.iii$, and comprise 3.30, 2.20, 1.80 eV and 3.50, 2.35, 1.85 eV, respectively. In addition to these three similarly looking molecular states, two additional states residing at lower binding energies are found in CoPyr on Ag(110): *cp.iv* at 1.05 eV and *cp.v* at 0.70 eV binding energies. These states are absent in case of Pyr on Ag(110) and their PADs reveal different features than those present in the first three molecular states. We note that they are present not only in the DFT data for CoPyr on Ag(110), but also in the free space DFT simulation of CoPyr, as can be seen by comparing Fig. 7.8(B.I) with Fig. 7.8(C.I). Because these two additional states are not present in Pyr data, we classify them as those arisen due to the presence of Co atom.

8

TIME-RESOLVED ARPES ON PENTACENE/AG(110). OUTLOOK

Up to now, orbital tomography has been employed only for the static reconstruction of molecular orbitals. However, in addition to spatial visualisation of stationary orbital distributions, it would be extremely interesting to map the evolution of excited state molecular orbitals at different moments in time. We make one step towards this goal by studying population dynamics in a bilayer of pentacene molecules self-assembled on a Ag(110) substrate. The first layer of pentacene serves as an electronic decoupling layer, to prevent the population of the LUMO of the second layer by substrate electrons. The goal is to resonantly excite the molecules in the second layer from HOMO to LUMO and probe the excited state with photons of high energy. Ideally, the probe light should have photon energies in the range of 25-35 eV, in order to cover the range of the photoelectron momenta between 1...2 Å^{-1}, containing most of the information about spatial distribution of orbitals. In principle, such combination is a prerequisite for time-resolved mapping of molecular electron dynamics, provided the PADs are acquired at a broad range of polar and azimuthal electron emission angles.

This chapter presents results obtained during the beamtime at the ARTEMIS beamline of the Central Laser Facility at the Rutherford Appleton Laboratory in the United Kingdom. The ARTEMIS beamline provides an exclusive combination of an Optical Parametric Amplifier, tuneable in a broad IR-UV spectral range, and a HHG laser source, essential for such kind of experiments.

8.1 PREPARATORY WORK AT UZH

Prior to the beamtime at ARTEMIS, preparation and characterisation of the samples was done at the UZH lab.

8.1.1 *Sample preparation and characterisation*

The samples were prepared in situ by evaporating pentacene molecules from a Knudsen cell onto the Ag(110) single crystal surface. The Ag(110) surface was cleaned by 2-3 cycles of sputter-anneal prior to the deposition of molecules. Sputtering was done with Ar ions at $3{\cdot}10^{-6}$ mbar pressure, 500 eV bias, $\approx$ 540 eV discharge voltage, 4.4 mA discharge current. The sputter current was $\approx 2\ \mu$A. The molecules were evaporated onto the Ag(110) surface held at room temperature. A monolayer was typically obtained after 12 min of evaporation at 200°C, a bilayer after 25 min. Fig. 8.1 shows the corresponding LEED patterns. The lattice vectors of the first layer of pentacene are shown in blue, while those of the second in red. The real space lattice vectors of the Ag(110) substrate, $\boldsymbol{a}_1 = 2.889$ Å and $\boldsymbol{a}_2 = 4.086$ Å, can be related to the lattice vectors of the monolayer, $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$, and of the second layer of the bilayer, $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$, via epitaxial matrices as specified in Table 8.1.

	RECIPROCAL SPACE	REAL SPACE
Monolayer	$\begin{pmatrix} \boldsymbol{b}_1^* \\ \boldsymbol{b}_2^* \end{pmatrix} = \dfrac{1}{11} \begin{pmatrix} 3.0 & -2.0 \\ 1.0 & 3.0 \end{pmatrix} \begin{pmatrix} \boldsymbol{a}_1^* \\ \boldsymbol{a}_2^* \end{pmatrix}$	$\begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{pmatrix} = \begin{pmatrix} 3.0 & -1.0 \\ 2.0 & 3.0 \end{pmatrix} \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{pmatrix}$
Second layer of bilayer	$\begin{pmatrix} \boldsymbol{c}_1^* \\ \boldsymbol{c}_2^* \end{pmatrix} = \dfrac{1}{11} \begin{pmatrix} 3.6 & -2.6 \\ 1.0 & 3.0 \end{pmatrix} \begin{pmatrix} \boldsymbol{a}_1^* \\ \boldsymbol{a}_2^* \end{pmatrix}$	$\begin{pmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \end{pmatrix} = \begin{pmatrix} 3\frac{5}{3} & -1\frac{5}{3} \\ 2.0 & 3.0 \end{pmatrix} \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{pmatrix}$

Table 8.1: Epitaxial matrices of real and reciprocal space lattices of pentacene on Ag(110) for monolayer and the second layer of bilayer. Substrate vectors are denoted by $\boldsymbol{a}_i$, those of the monolayer by $\boldsymbol{b}_i$, of the second layer of bilayer by $\boldsymbol{c}_i$. The superscript * denotes reciprocal space vectors.

Ultraviolet photoelectron spectroscopy (UPS) was used to characterise the valence orbital structure of pentacene to determine the energy position of molecular states. Fig. 8.2(A) shows spectra recorded from the bare Ag(110), a mono- and a bilayer of pentacene on Ag(110) using He I light. While the monolayer data reveal only two peaks, which we assign to the HOMO and LUMO, the bilayer spectrum has three

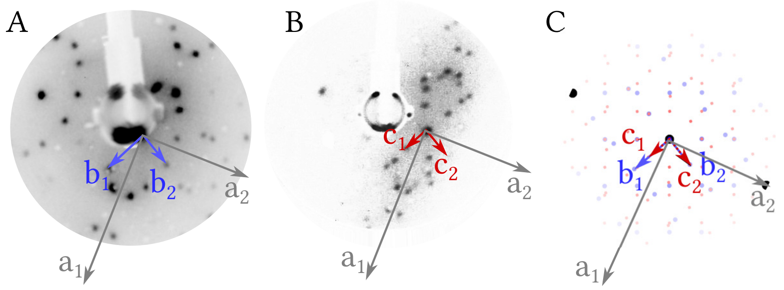


Figure 8.1: LEED patterns recorded from pentacene on Ag(110) at 31 eV electron energy. (A) – Monolayer. (B) – Bilayer. (C) – Overlaid reciprocal lattices of of the first (blue) and of the second (red) layer together with the corresponding reciprocal lattice vectors. Reciprocal lattice vectors of the substrate are shown in grey.

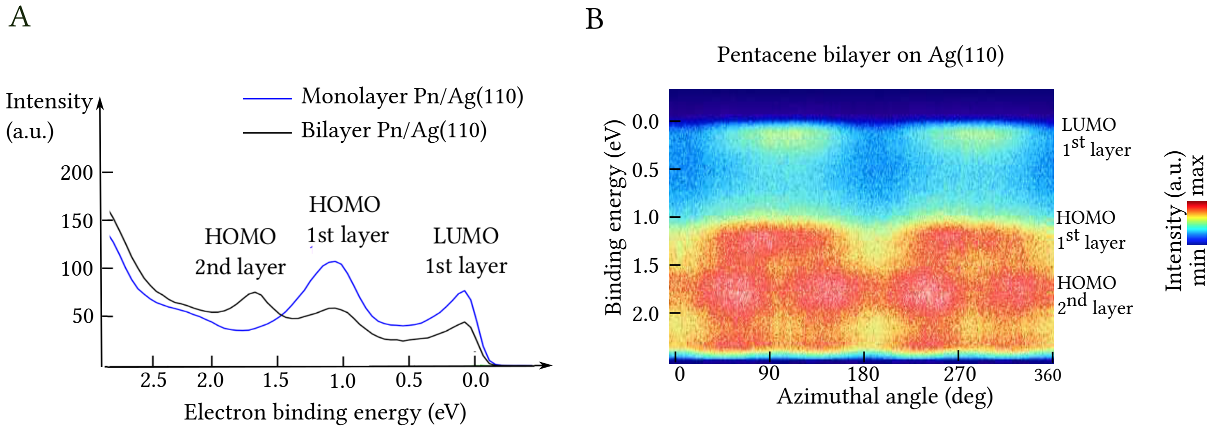


Figure 8.2: (A) – UPS recorded from a monolayer (blue) and bilayer (black) of pentacene on Ag(110) using He I light. (B) – The same data as in (A) for bilayer of pentacene, but recorded for the whole range of the azimuthal angles by rotating the sample around its normal at the polar angle of 50°.

molecular states at ≈ 0.1, ≈ 1.1 and ≈ 1.7 eV binding energy, corresponding to the LUMO of the first, HOMO of the first and HOMO of the second layer, respectively. In both cases, the presence of the LUMO peak in the spectra is due to the LUMO population via charge transfer from the Ag(110) substrate. Fig. 8.2(B) shows UPS data recorded from the bilayer of pentacene on Ag(110), in which the electron binding energy is plotted vs. the azimuthal angle of the sample. The symmetries of the respective molecular states can be resolved and after comparing with the literature data [3], the initial assignment of the peaks can indeed be confirmed.

8.1.2 *Temporal characterisation of laser pulses*

Apart from the static characterisation of the samples with LEED and UPS, another goal of the preparatory experiments at UZH was to determine the resonant wavelength of the HOMO-LUMO transition in the second layer of pentacene. The setup for this measurements is described in section 4. Due to the time limitation, we were able to perform only temporal characterisation of the laser pulses.

For 2PPE measurements, the pump and probe laser pulses must overlap in space and time. To check the spatial overlap, pump and probe beams were picked off in front of the view port of the vacuum chamber by a mirror and directed onto a CCD camera. The distance between the mirror and the CCD camera was equal to that between the mirror and the sample surface in the vacuum chamber. The temporal overlap was achieved by ensuring that both optical paths are (i) collinear, (ii) equal in their length within the error of a few millimetres and (iii) that the pump and probe spots overlap on the CCD camera over the whole length of the delay stage. For the purpose of finding the temporal overlap, we chose a GaAs (111) single crystal. Due to long living processes in the uncleaned GaAs crystal, time-zero, i.e. temporal overlap, can be found relatively simply, provided the pulses are in spatial overlap on the surface of the sample over the length of the delay stage. Fig. 8.3(A) shows a delay scan recorded from the surface of GaAs(111) with $\lambda = 530$ nm (≈ 2.34 eV) and $\lambda = 410$ nm (≈ 3.02 eV) laser pulses with the step size of 100 fs. We applied a bias voltage of -20 V and recorded the signal at 21 eV electron kinetic and 20 eV pass energies. In Fig. 8.3, the bias voltage was subtracted. The increase in photoemission signal around 1.8 eV kinetic energy marks the position of the temporal overlap of the pulses. The band gap of GaAs is 1.42 eV and the valence band maximum (VBM) lies around 5.5 eV below the vacuum level.

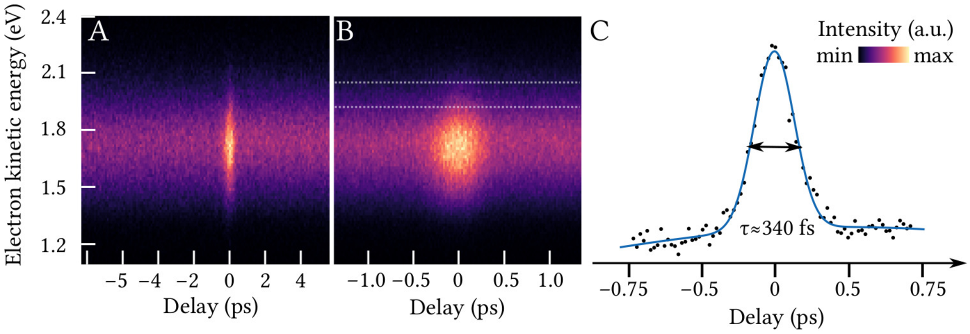


Figure 8.3: Delay scans on GaAs(111). Pump pulse: λ = 410 nm (≈ 3.02 eV), probe pulse: λ=530 nm (≈ 2.34 eV). (A) – Survey scan with the step size Δt = 100 fs. Pump and probe pulses are have approximately equal energy of about 23 nJ per pulse. (B) – Zoom into (A) with the step size Δt = 20 fs. (C) – Intensity distribution obtained by integration between the white dashed lines in (B). Fit with a Gaussian function results in the full-width-half-maximum of 340 fs.

Given the photon energies are equal to 2.34 eV (green) and 3.02 eV (blue), the origin of the signal can be thereby attributed to a number of different multi-photon photoemission processes shown in Fig. 8.4. The most probable ones are: (1) Simultaneous absorption of two blue photons → direct electron escape into vacuum; (2) Excitation of "hot" electrons via absorption of one blue photon → fast relaxation of the "hot" electrons to the bottom of the conduction band → absorption of two green photons → electron escape into vacuum; (3) Excitation of "hot" electrons via absorption of one green photon → fast relaxation of the "hot" electrons to the bottom of the conduction band → absorption of two green photons → electron escape into vacuum. A detailed analysis of these and other possible electron photoemission processes was done by M. Meier in his Bachelor thesis [164]. The shape of the signal around time-zero is determined by the cross-correlation of the pump and probe pulses and by the lifetime of the underlying relaxation processes. To unambiguously determine the width of the cross-correlation, it is necessary to know the latter. In this work, the precise determination of the lifetimes of the electronic processes was not possible, therefore only the upper limit for the width of the cross-correlation could be determined [164]. Fig. 8.3(B) is a zoomed version of (A) recorded at the same settings, but with the step size of 20 fs.

By integrating the signal between the white lines in Fig. 8.3(B) and normalising it to 1, we obtained a Gaussian distribution and fitted it with the Gaussian function $f(t) = e^{-\frac{t-t_0}{2\sigma^2}}$, where t_0 is time zero delay and σ is the standard deviation of the distribution, related to the full width at half maximum (FWHM) as FWHM $= 2\sqrt{2\ln 2}\sigma$. From the fit shown in Fig. 8.3(B), we obtained the FWHM of $\tau \approx 340$ fs. Obviously,

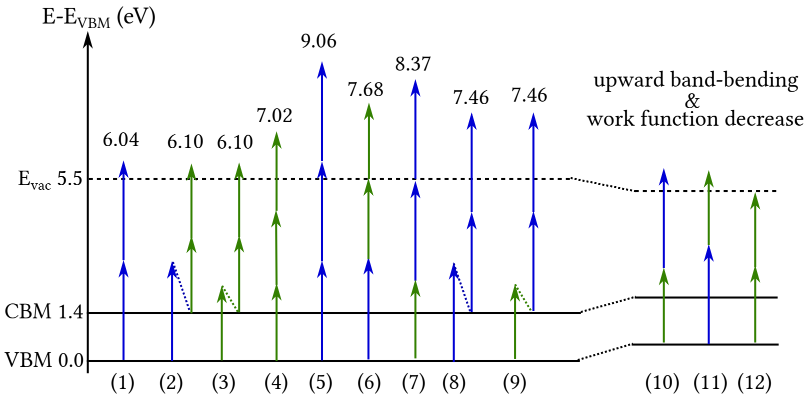


Figure 8.4: Possible multi-photon photoemission processes with final state energies. Blue and green arrows denote possible transitions due to absorption of 410 nm (≈ 3.02 eV) and 530 nm (≈ 2.34 eV) photons, respectively. CBM stands for conduction band minimum. Adapted from [164].

this is a very crude estimate of the width of the cross-correlation. A more precise model, including an exponential function, to account for the lifetimes of the involved electronic processes, was developed by M. Meier in his Bachelor thesis [164].

8.2 ARTEMIS

The measurements at the ARTEMIS beamline aimed at excitation of the HOMO-LUMO transition in the second layer of the bilayer of pentacene on Ag(110) and probing of the excited state dynamics with the HHG photons. The samples consisted of a bilayer of pentacene on Ag(110) and were prepared and characterised with LEED and UPS as described in the previous section.

The resonant wavelength for the HOMO-LUMO transition was not known at the moment of measurements. In literature, different values were reported for different methods. Dougherty et al. [165] reported a transport gap of 3.6±0.2 eV measured with STM on a bilayer of pentacene. Amy et al. [166] used a combination of photoelectron spectroscopy and inverse photoelectron spectroscopy to study multilayers of pentacene and reported 2.88 eV energy for the HOMO-LUMO transition in 25 Å thick films. Han et al. employed near-ultraviolet inverse photoelectron spectroscopy and obtained the band-gap energy of 10 nm pentacene film to be 2.20±0.6 eV, consistent with Amy et al. [166] data. Repp and Meyer [167] studied a monolayer of pentacene decoupled from Cu(111) by a bilayer of a NaCl film. By STM, they found a band gap of 302 nm (4.1 eV).

We chose a pump wavelength of 420 nm (2.95 eV), as it was on the order of values reported in literature. The photon energy of probe light was selected to be 25.5 eV (17th harmonic of the HHG radiation). Time zero was estimated based on spectral shifts due to space charge effects [168, 169]. Though, for measurements, we set the intensity of the HHG light to moderately low values in order to keep the space charge effects to a minimum. Prior to acquisition of the PADs covering the whole hemisphere, it was necessary to confirm the suitability of the chosen pump wavelength for the excitation of the HOMO-LUMO transition. This was done by recording time-resolved data only at selected positions in reciprocal space, i.e. at characteristic features of the frontier orbitals, HOMO and LUMO, at around 1.0...1.5 Å^{-1}. For this purpose, we aligned the sample position without the pump light, but using the probe pulses only. We optimised an azimuthal position of the sample to maximise the signal. Maximal counts rate was obtained at 50° polar angle, which corresponded to about 1.5 Å^{-1} in reciprocal space at 25.5 eV photon energy, as expected for the position of HOMO and LUMO features from literature [3] and He I experimental data.

The UPS data recorded with He I light, shown in Fig. 8.2, show that the occupied molecular states are separated from each other by 0.5 to 1.0 eV and their FWHM is on the order of 200 meV. Because the FWHM of HHG light was around 900 meV, it was challenging to resolve the individual HOMO and LUMO peaks in the recorded spectrum. Moreover, we found that the chosen 420 nm wavelength for the pump did not lead to any excitation, as there were no visible changes in the recorded data. The variations in the signal intensity around time zero were on the order of noise level and we were not able to resolve any increase in the signal above the Fermi level, as it was expected. Hence, in order to boost the signal to noise ratio and, possibly, provide means for the resonant excitation of the HOMO-LUMO transition, we prepared a multilayer of pentacene on Ag(110). Using the same settings, i.e. 420 nm pump and 25.5 eV probe laser light and 50° polar angle, we acquired the data during $\approx$ 12 hours to achieve high statistics.

Fig. 8.5 shows CCD snapshots recorded at different moments of time. The Fermi level is located at $\approx$ 21 eV. The data recorded at -2 ps, i.e. far from the expected position of the time zero, are shown in Fig. 8.5(A,B). (A) are the raw data and (B) are the same data, but after subtraction of the signal at -2.7 ps. As expected for a signal distribution far away from time-zero, any substantial changes in the signal intensity, which can be attributed to excited dynamics, cannot be identified.

In contrast, Fig. 8.5 (C,D) shows the same datasets as in (A,B), but in the vicinity of the time zero. (C) are the raw data and (D) are the same data, but after subtraction of the signal at -2.7 ps. An increase of the signal intensity distribution on the order of a few percent is seen around the Fermi level, while the signal distribution at $\approx$ 20.5 eV and 18.5 eV are depleted by the same amount.

To study this effect in more detail, we integrated the ARPES data shown in Fig. 8.5 over the whole range of angles at each time delay. This procedure yielded energy distribution curves, which were plotted as a function of time delay. At each time step, energy distribution curves were normalised by the corresponding maximal intensity values to increase signal-to-noise ratio. An increase in the signal intensity is seen around 0 ps at electron kinetic energies at ≈ 19 eV and between 20.5 and 22.5 eV. While it is difficult to interpret this effect, it can certainly be attributed to some laser induced electron excitation in the sample. In fact, our result is in agreement with that obtained by Amy et al. [166], who reported 2.88 eV energy for the HOMO-LUMO transition in 25 Å thick pentacene films. We were not able to exactly determine the thickness of the pentacene multilayer, but it was certainly on that order.

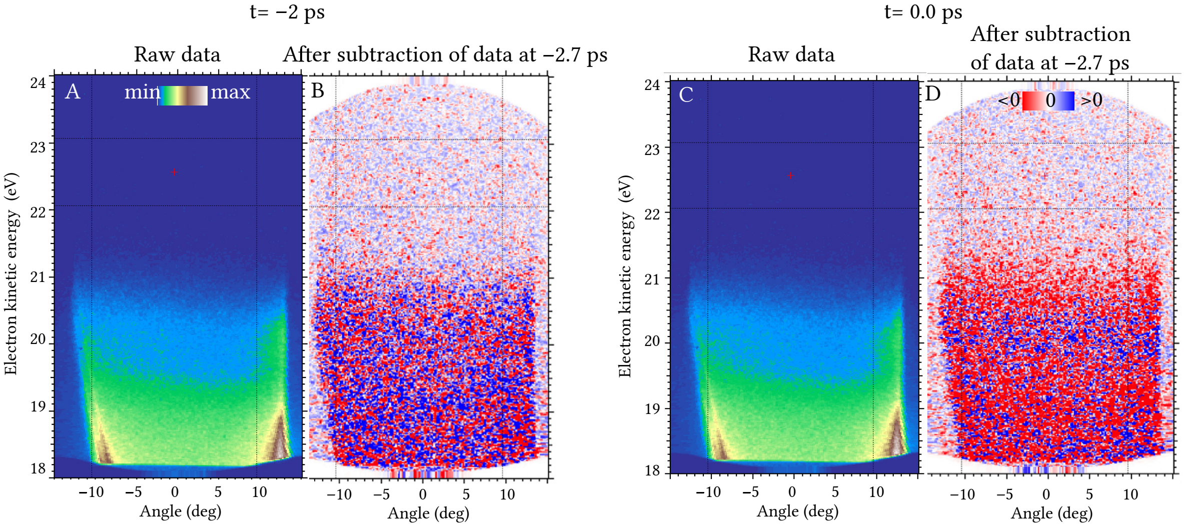


Figure 8.5: ARPES data recorded with 420 nm pump and 25.5 eV probe laser light from a multilayer of Pentacene on Ag(110). (A, B) – Data recorded at −2 ps before the expected location of the time zero. (C, D) – Data recorded around 0 ps time delay. (A, C) Raw data. (B, D) After subtraction of the corresponding data recorded at −2.7 ps.

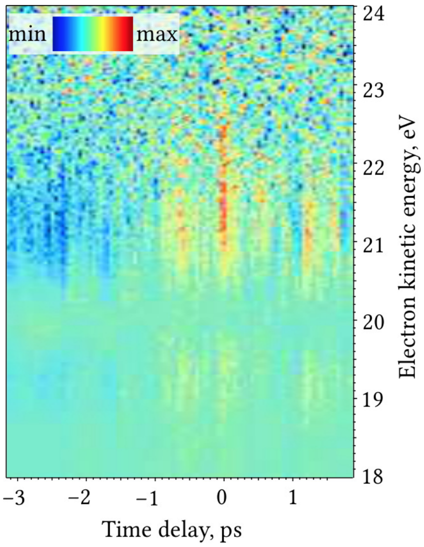


Figure 8.6: Delay scans recorded with 420 nm pump and 25.5 eV probe high harmonic light from a multilayer of Pentacene on Ag(110), normalised at each time step by the corresponding maximal intensity value.

This result provides a good starting point for more systematic and rigorous studies of electron dynamics in pentacene adlayers. Certainly, time-resolved orbital tomography of the second layer of the pentacene bilayer and studies on population dynamics require exact knowledge about the resonant excitation wavelength for the HOMO-LUMO transition. Only with such knowledge, one will be able to study the electron population dynamic and then extend the established methods towards time-resolved mapping of the photoelectron angular distributions and reconstruction of the corresponding wave functions at different moments in time.

SUMMARY AND OUTLOOK

In this work, we advanced orbital tomography both theoretically and experimentally. We proposed solving the phase problem in orbital tomography in a robust manner, by finding its analogue in optics. We demonstrated it by conducting an optical coherent diffraction imaging experiment, in which the amplitude and phase distributions of micrometer-sized structures were reconstructed solely from their diffraction patterns by employing state-of-the-art iterative phase retrieval algorithms: the shrinkwrap algorithm of Marchesini [12], error reduction algorithm of Fienup [11], and the modified version of the hybrid input-output algorithm of Fienup [11], which was proposed by Harder [13] for reconstruction of complex-valued object distributions. By applying these algorithms to angle-resolved photoemission data recorded from the lowest unoccupied molecular orbital of pentacene on Ag(110), we were able to reconstruct its amplitude and phase distribution without symmetrisation of the experimental data and assumptions about shape of the orbital distribution and its symmetry properties.

To facilitate subtraction of the background features from the experimental data prior to their use in phase retrieval, we developed an efficient image processing procedure, which is based on the maximisation of mutual information algorithm of Mattes [14]. We employed this procedure to process experimental angle-resolved photoemission data recorded from the highest occupied molecular orbital of pentacene on Ag(110). Moreover, we improved the quality of reconstructions with phase retrieval by dividing the reconstruction process into two parts. Conventionally, the reconstruction is performed using the shrinkwrap algorithm. However, the algorithms often stagnate, delivering artefacts instead of correct reconstructed object distributions. We solved this issue by introducing a second reconstruction run, in which the centred and thresholded reconstructed object distribution obtained after the first reconstruction run was used as a tight object support. This approach prevented the algorithm from stagnation and, in addition, eliminated the ambiguity about the location of the geometrical centre of the orbital distribution. We demonstrated the feasibility of the proposed approach by successfully reconstructing the amplitude and phase distribution of the highest occupied molecular orbital of pentacene. The results of the reconstruction were compared with density functional theory simulations and revealed good agreement. We provide a detailed procedure for processing of the density functional theory data in order to provide their quantitative comparison with the experimental results and reconstructed orbital distributions. We discovered that the 2D orbital distributions, reconstructed by phase retrieval, can be viewed as a superposition of the “in-focus” orbital distribution at the plane $z = 0$ and “out-of-focus” orbital distributions at other planes $z = \text{const}$, whose strength and blur are set by experimental conditions. Additionally, we came to the conclusion that one can reconstruct full 3D orbital distribution solely from a single set of the complex-valued 2D distributions in reciprocal space, the phase of which was obtained in the phase retrieval procedure. However, for this purpose, substantial improvement of the axial resolution by increase of the numerical aperture of the system or photon energies is required. Unambiguous 3D reconstruction requires photoemission data acquired at several photon energies though.

Experimentally, we extended orbital tomography to a new class of molecules: pyrphyrin and cobalt pyrphyrin on Ag(110). We were able to record photoelectron angular distributions from 3 molecular states of pyrphyrin and 5 molecular states of cobalt pyrphyrin. Thorough analysis of the low energy electron diffraction data and the simulated photoelectron angular distributions obtained from the density functional theory and independent-atomic-centre approximation, indicates that the molecules order themselves on the surface of Ag(110) in three and two rotational domains, respectively. In spite of this fact, we used the experimental data to reconstruct real space distributions. Owing to the presence of multiple rotational domains, these can no longer be interpreted as orbital distributions, albeit they share some of the features with the 2D orbital distributions obtained from the density functional theory. The only reason for the similarity is that the data share some of the features in reciprocal space, which make them look similar in real space as well. We would like to highlight that in case when multiple rotational domains of molecules result in incoherent summation of the individual photoelectron distributions in the detector plane, reconstructed distributions in real space loose their connection to true orbital distributions and, strictly speaking, represent only some fictive real space distributions that have no analogue in the real world.

Reconstruction of orbital distributions with phase retrieval algorithms requires data acquired from a single rotational domain of molecules only and, ideally, in the absence of spurious background features. The reason for this second requirement is minimal user intrusion into raw data for phase retrieval algorithms, which ensures unbiasedness of reconstructions. Additionally, to ensure stability of the reconstruction procedure and excellent quality of reconstructed data, experimental data must be acquired with high statistics. This requirement is particularly important if the goal is time-resolved reconstruction of orbital distributions.

Finally, we conducted a proof-of-principle time-resolved experiment on a multilayer of pentacene on the Ag(110). In fact, the goal was to excite the HOMO-LUMO transition in the second layer of bilayer pentacene and then probe the electron population dynamics with photons generated through high harmonic generation. Instead, we excited a multilayer of pentacene. By exciting the sample with 2.95 eV pump and 25.5 eV probe photons, we observed an increase of photoelectron signal around the Fermi level followed by depletion 0.5 and 2.5 eV below. We consider this result to be a good starting point for more systematic and rigorous studies of electron dynamics in pentacene adlayers. For successful realisation of such experiments, exact knowledge about the resonant excitation wavelength for the HOMO-LUMO transition is required. In this case, one will be able to study the electron population dynamics and then extend the established methods towards time-resolved mapping of the photoelectron angular distributions and reconstruction of the corresponding wave functions at different moments in time.

X-RAY PHOTOELECTRON DIFFRACTION

Spatial orientation of the individual CoPyr and Pyr molecules with respect to the Ag(110) substrate and the apparent existence of multiple rotational domains were to be determined by x-ray photoelectron diffraction (XPD) [155]. The measurements were done at the PEARL beamline of the Swiss Light Source at Paul Scherrer Institute in Villigen (Switzerland). The strongest signal was expected close to 90° polar angle, in view of the flat geometry of the molecules.

PREPARATION AND CHARACTERISATION OF THE SAMPLE The Ag(110) substrate was prepared using standard sputter-anneal cleaning cycles. Though immediately after the cleaning of the surface, the sample quality was checked with x-ray photoelectron spectroscopy (XPS) and did not reveal any contaminants, LEED showed an unknown superstructure. The XPS measurement showed the presence of a C1s peak. The subsequent sputter-anneal cycles did not improve the quality of the LEED patterns. In the course of the experiment, we noticed that the unknown superstructure was not present if the sample temperature was above 125°C. The strongest pattern we observed was the one at ≈ 25°C. Fig. 8.7 shows LEED patterns acquired at different temperatures of the sample. The process was reversible, i.e. the superstructure disappeared whenever the sample was heated above 125°C and then reappeared when the sample was cooled down to room temperature.

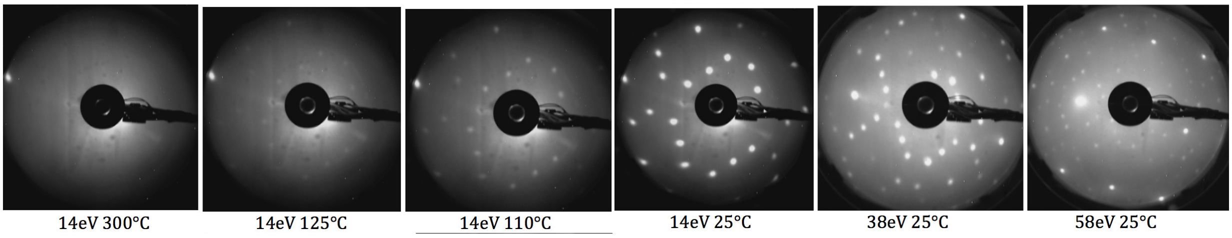

14eV 300°C
14eV 125°C
14eV 110°C
14eV 25°C
38eV 25°C
58eV 25°C

Figure 8.7: LEED patterns recorded from bare Ag(110) substrate at PEARL beamline

To clarify the nature of these unknown features in LEED, we repeated the same cleaning procedures with the same crystal at the UZH lab and were not able to find any traces of the unknown superstructure. The surface of the crystal healed itself to its normal state. Thus, we can attribute the superstructure encountered during the beamtime to contaminants present in the chamber and not to an unknown form of the Ag(110) surface reconstruction.

EVAPORATION OF MOLECULES AND ACQUISITION OF THE XPD PATTERN Despite of the presence of the unknown superstructure, we evaporated CoPyr molecules onto the surface of the Ag(110) substrate in its deteriorated state.

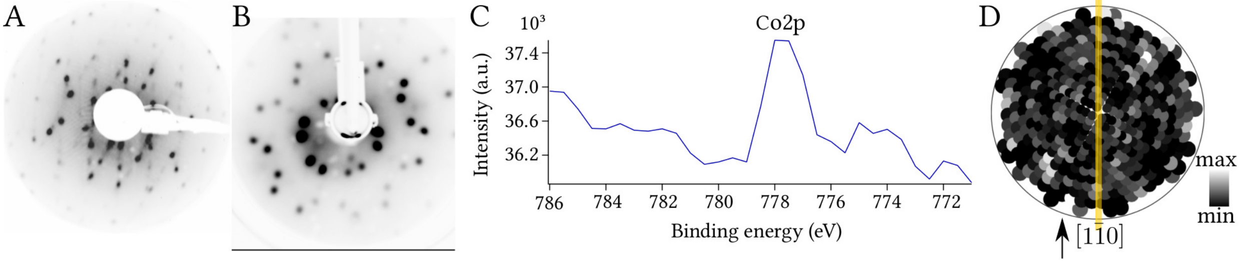


Figure 8.8: (A,B) – LEED patterns acquired from CoPyr on Ag(110) at 29 eV electron energy. (A) – Pattern acquired at PEARL. (B) – Reference pattern acquired at UZH. (C, D) – XPS (C) and XPD (D) acquired with 1000 eV photon energy, 50 eV pass energy in Co2p region. In (D), the yellow line denotes the $[1\bar{1}0]$ high symmetry direction of the substrate.

We performed evaporation of CoPyr at 330°C for 20 min, which should have resulted in $\approx$ 1 ML. Fig. 8.8(A) shows the resulting LEED image. Fig. 8.8(B) shows the reference image recorded earlier in our lab. The LEED pattern acquired during the beamtime mostly contains the features of the unknown superstructure.

Despite the poor quality of the LEED data, we performed XPD measurements. Fig. 8.8(C) shows the XPS survey spectrum acquired in the region of the Co 2p peak at 1000 eV photon energy and 50 eV pass energy. Fig. 8.8(D) shows the respective XPD pattern at the binding energy of the Co 2p3/2 peak at 778 eV. The yellow line highlights the $[1\bar{1}0]$ high symmetry direction of the substrate. The angle between the high symmetry direction and the axis through two intensity maxima, located at about 84° polar angle, amounts to 42°. These maxima might be attributed to scattering of the electron waves emerging from the cobalt atom, located in the centre of the molecule, on the neighbouring nitrogen atoms. Though, this claim, due to the lack of better statistics, definitely needs further investigations.

BIBLIOGRAPHY

1. Puschnig, P., Berkebile, S., Fleming, A. J., Koller, G., Emtsev, K., Seyller, T., Riley, J. D., Ambrosch-Draxl, C., Netzer, F. P. & Ramsey, M. G. Reconstruction of molecular orbital densities from photoemission data. *Science* **326,** 702–6 (2009).
2. Puschnig, P., Koller, G., Draxl, C. & Ramsey, M. G. *Small Organic Molecules on Surfaces* (ed Sitter, H.) 3–23 (Springer, 2013).
3. Lüftner, D., Ules, T., M., E., Koller, G., Soubatch, S., Tautz, F. S., Ramsey, M. G. & Puschnig, P. Imaging the wave functions of adsorbed molecules. *PNAS* **111,** 605–610 (2014).
4. Weiss, S. *et al.* Exploring three-dimensional orbital imaging with energy-dependent photoemission tomography. *Nature Communications* **6** (2015).
5. Kliuiev, P., Latychevskaia, T., Osterwalder, J., Hengsberger, M. & Castiglioni, L. Application of iterative phase-retrieval algorithms to ARPES orbital tomography. *New J. Phys.* **18,** 093041 (2016).
6. Goldberg, S. M., Fadley, C. S. & Kono, S. Photoelectric cross-sections for fixed-orientation atomic orbitals: Relationship to the plane-wave final state approximation and angle-resolved photoemission. *Solid State Communications* **28,** 459–463 (1978).
7. Wiessner, M., Hauschild, D., Sauer, C., Feyer, V., Schoell, A. & Reinert, F. Complete determination of molecular orbitals by measurement of phase symmetry and electron density. *Nature Communications* **5,** 4156 (2014).
8. Miao, J. Extending the methodology of X-Ray crystallography to allow imaging of micrometer-sized non-crystalline specimens. *Nature* **400,** 342–344 (1999).
9. Miao, J., Sayre, D. & Chapman, H. N. Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects. *J. Opt. Soc. Am. A* **15,** 1662–9 (1998).
10. Miao, J., Ishikawa, T., Anderson, E. H. & Hodgson, K. O. Phase retrieval of diffraction patterns from noncrystalline samples using the oversampling method. *Physical Review B* **67,** 174104 (2003).
11. Fienup, J. R. Reconstruction of an object from the modulus of its Fourier transform. *Optics Letters* **3,** 27–29 (1978).
12. Marchesini, S., He, H., Chapman, H. N., Hau-Riege, S. P., Noy, A., Howells, M. R., Weierstall, U. & Spence, J. C. H. X-Ray image reconstruction from a diffraction pattern alone. *Physical Review B* **68,** 140101 (2003).
13. Harder, R., Liang, M., Sun, Y., Xia, Y. & Robinson, I. K. Imaging of complex density in silver nanocubes by coherent x-ray diffraction. *New J. Phys.* **12,** 035019 (2010).
14. Mattes, D., Haynor, D. R., Vessele, H., Lewellen, T. K. & Eubank, W. *Nonrigid multimodality image registration* in *Medical Imaging 2001: Image Processing* (eds Sonka, M. & Hanson, K. M.) **4322** (2001), 1609–1620.
15. Zitova, B. & Flusser, J. Image registration methods: a survey. *Image and Vision Computing* **21,** 977–1000 (2003).
16. Hohenberg, P. & Kohn, W. Inhomogeneous Electron Gas. *Physics Review* **136,** 864–871 (1964).
17. Kohn, W. & Sham, L. J. Self-Consistent Equations Including Exchange and Correlations Effects. *Physical Review* **140,** 1133–1138 (1965).
18. Barton, J. Photoelectron Holography. *Phys. Rev. Lett.* **61,** 1356–1359 (1988).
19. Joliat, E., Schnidrig, S., Probst, B., Bachmann, C., Spingler, B., Baldridge, K. K., von Rohr, F., Schillingb, A. & Alberto, R. Cobalt complexes of tetradentate, bipyridine- based macrocycles: their structures, properties and photocatalytic proton reduction. *Dalton Transactions* **45,** 1737–1745 (2016).
20. Gurdal, Y., Luber, S., Hutter, J. & Iannuzzi, M. Non-innocent adsorption of Co-pyrphyrin on rutile(110). *Phys. Chem. Chem. Phys.* **17,** 22846 (2015).
21. Hertz, H. Ueber einen Einfluss des ultravioletten Lichtes auf die electrische Entladung. *Ann. Physik* **33,** 983–1000 (1887).

22. Thomson, J. On the masses of the ions in gases at low pressures. *Philos. Mag.* **48,** 547 (1899).
23. Einstein, A. Ueber einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Ann. Physik* **17,** 132–148 (1905).
24. Berglund, C. & Spicer, W. Photoemission Studies of Copper and Silver: Theory. *Phys. Rev. A* **136,** 1030–1044 (1964).
25. Huefner, S. *Photoelectron spectroscopy: Principles and Applications* (Springer, 2003).
26. Osterwalder, J. *Surface and Interface Science* (ed Wandelt, K.) chap. Photoelectron Spectroscopy and Diffraction (Wiley-VCH, 2014).
27. Feibelman, P. & Eastman, D. Photoemission spectroscopy - Correspondence between quantum theory and experimental phenomenology. *Phys. Rev. B* **10,** 4932–4947 (1974).
28. Landau, L. & Lifshitz, E. *Quantum Mechanics. Non-Relativistic Theory (in Russian)* (Nauka, Moscow, 1989).
29. Levich, V. *Course of Theoretical Physics (in Russian)* (Nauka, 1971).
30. Gadzuk, J. W. Angular distributions of electrons photoemitted from chemisorbed atoms. *Solid State Communications* **15,** 1011–1016 (1974).
31. Gerchberg, R. W. & Saxton, W. O. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik* **35,** 273–246 (1972).
32. Permien, T., Engelhardt, R. & Feldmann, C. A. Angle-resolved photoemission from oriented films of lead phthalocyanine on a Cu(100) surface. *Chem. Phys. Lett.* **98,** 527–530 (1983).
33. Richardson, N. Comments on angle-resolved photoemission from oriented films of lead phthalocyanine on a Cu(100) surface. *Chem. Phys. Lett.* **102,** 390–391 (1983).
34. Grobman, W. D. Angle-resolved photoemission from molecules in the independent-atomic-center approximation. *Phys. Rev. B* **17,** 4573–4585 (1978).
35. Gross, H. *Handbook of Optical Systems. Fundamentals of Technical Optics* (ed Gross, H.) (Wiley-VCH, 2005).
36. Green, H. & Wolf, E. A Scalar Representation of Electromagnetic Fields. *Proc. Phys. Soc. Sec. A* **66** (1953).
37. Singer, W., Totzek, M. & Gross, H. *Handbook of Optical Systems. Physical Image Formation* (ed Gross, H.) (Wiley-VCH, 2005).
38. Ratcliffe, J. A. Some aspects of the diffraction theory and their application to the ionosphere. *Reports on Progress in Physics* **19,** 188 (1956).
39. Goodman, J. W. *Introduction to Fourier Optics* 2nd ed. (The McGraw-Hill Companies, Inc., 1996).
40. Dainty, J. & Fienup, J. *Image recovery: theory and applications* (ed Stark, H.) (Elsevier, 1987).
41. Röntgen, W. Ueber eine neue Art von Strahlen. *Sitzungsberichte der Physikalisch-medizinischen Gesellschaft zu Wuerzburg,* 137–147 (1895).
42. Bravais, A. *Mémoire sur les systèmes formés par les points distribués régulièrement sur un plan ou dans l'espace* (J. Ecole Polytech, 1850).
43. Sommerfeld, A. Ueber die Beugung der Roentgenstrahlen. *Annalen der Physik,* 473–506 (1912).
44. Friedrich, W., Knipping, P & Laue, M. Interferenz-Erscheinungen bei Röntgenstrahlen. *Annalen der Physik* **346,** 971–988 (1913).
45. Sayre, D. Some implications of a theorem due to Shannon. *Acta Cryst.* **5,** 843 (1952).
46. Shannon, C. E. Communication in the presence of noise. *Proceedings of the Inst. Radio Engrs.* **37,** 10–21 (1949).
47. Fienup, J. R., Crimmins, T. R. & Holsztynski, W. Reconstruction of the support of an object from the support of its autocorrelation. *J. Opt. Soc. Am.* **72,** 610–624 (1982).
48. J.R.Fienup. Reconstruction of a complex-valued object from the modulus of its Fourier transform using a support constraint. *J. Opt. Soc. Am. A* **4,** 118–123 (1987).
49. Fienup, J. Phase retreival algorithms: a comparison. *Applied Optics* **21** (1982).

50. Williams, G. J., Pfeifer, M. A., Vartanyants, I. A. & Robinson, I. K. *Physical Review B* **73,** 094112 (2006).

51. Bates, R. & Tan, D. *Fourier phase retrieval when the image is complex* in *Inverse Optics II* (eds Devaney, A. & Bates, R.) **558** (1985), 54–59.

52. Weierstall, U., Q.Chen, Spence, J., Howells, M., Isaacson, M. & Panepucci, R. Image reconstruction from electron and X-ray diffraction patterns using iterative algorithms: experiment and simulation. *Ultramicroscopy* **90,** 171–195 (2002).

53. Fienup, J. & Wackerman, C. Phase-retrieval stagnation problems and solutions. *Opt. Soc. Am. A* **3,** 1897–1907 (1986).

54. Chapman, H., Barty, A., Marchesini, S., Noy, A. & Hau-Riege, S. High-resolution ab initio three-dimensional x-ray diffraction microscopy. *J. Opt. Soc. Am. A* **23,** 1179 (2006).

55. Vartanyants, I. & Robinson, I. Partial coherence effects on the imaging of small crystals using coherent x-ray diffraction. *Journal of Physics: Condensed Matter* **13,** 10593–10611 (2001).

56. Williams, G., Pfeifer, M., Vartanyants, I. & Robinson, I. Three-dimensional imaging of microstructures in Au nanocrystals. *Phys. Rev. Lett.* **90,** 175501 (2003).

57. Newton, M., Harder, R., Huang, X., Xiong, G. & Robinson, I. Phase retrieval of diffraction from highly strained crystals. *Phys. Rev. B* **82,** 165436 (2010).

58. Latychevskaia, T., Chushkin, Y., Zontone, F. & Fink, H.-W. Imaging outside the box: Resolution enhancement in X-ray coherent diffraction imaging by extrapolation of diffraction patterns. *Applied Physics Letters* **107,** 183102 (2015).

59. Shapiro, D. *et al.* Biological imaging by soft x-ray diffraction microscopy. *PNAS* **102,** 15343–15346 (2005).

60. Thibault, P., Elser, V., Jacobsen, C., Shapiro, D. & Sayre, D. Reconstruction of a yeast cell from X-ray diffraction data. *Acta Cryst.* **A62,** 248–261 (2006).

61. Latychevskaia, T. & Fink, H.-W. Coherent microscopy at resolution beyond diffraction limit using post-experimental data extrapolation. *Applied Physics Letters* **103,** 204105 (2013).

62. Brigham, E. *The Fast Fourier Transform and its applicaitons* (Pearson, 1988).

63. Latychevskaia, T. & Fink, H.-W. Practical algorithms for simulation and reconstruction of digital in-line holograms. *Appl. Opt.* **54,** 2424–2434 (2015).

64. Cooley, J. & Tukey, J. An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation* **19,** 297–301 (1965).

65. Bruck, Y. M. & Soldin, L. G. On the ambiguity of the image reconstruction problem. *Optics Communications* **30,** 304–307 (1979).

66. Damascelli, A. Probing the electronic structure of complex systems by ARPES. *Physica Scripta* **T109,** 61–74 (2004).

67. Wiemann, C., Patt, M., Krug, I. P., Weber, N. B., Escher, M., Merkel, M. & Schneider, C. M. A New Nanospectroscopy Tool with Synchrotron Radiation: NanoESCA at Elettra. *Surf. Sci. Nanotech* **9** (2011).

68. Turcu, I., Springate, E. & Froud, C. *Ultrafast science and development at the ARTEMIS facility* in *ROMOPTO 2009: Ninth conference on optics: micro- to nanophotonics II* (ed Vlad, V.) **7469** (2009), 746902–1.

69. Greif, M., Castiglioni, L., Becker-Koch, D., Osterwalder, J. & Hengsberger, M. Acquisition of photoelectron diffraction patterns with a two-dimensional wide-angle electron analyzer. *Journal of Electron Spectroscopy and Related Phenomena* **197,** 30–36 (2014).

70. Greif, M. *Time-resolved photoelectron diffraction* PhD thesis (University of Zurich, 2014).

71. Schneider, C. *et al.* Expanding the view into complex material systems: From micro-ARPES to nanoscale HAXPES. *Journal of Electron Spectroscopy and Related Phenomena* **185,** 330 (2012).

72. Tusche, C., Krasyuk, A. & Kirschner, J. Spin resolved bandstructure imaging with a high resolution momentum microscope. *Ultramicroscopy* **159,** 520–529 (2015).

73. Ogawa, S., Narushima, R. & Arai, Y. Aza macrocycle that selectively binds lithium ion with color change. *J. Am. Chem. Soc.* **106** (1984).
74. Mette, G., Sutter, D., Gurdal, Y., Schnidrig, S., Probst, B., Ianuzzi, M., Hutter, J., Alberto, R. & Osterwalder, J. From porphyrins to pyrphyrins: adsorption study and metalation of a molecular catalyst on Au(111). *Nanoscale* **8,** 7958 (2016).
75. De Broglie, L. *Recherches sur la theorie des quanta* (1925).
76. De Broglie, L. *Nobel Lecture: The Wave Nature of the Electron* <`http://www.nobelprize.org/nobel_prizes/physics/laureates/1929/broglie-lecture.html`>.
77. Davisson, C. & Kunsman, C. H. The scattering of electrons by nickel. *Science* **LIV,** 522–524 (1921).
78. Davisson, C. & Germer, L. The scattering of electrons by a single crystal of nickel. *Nature* **119,** 558–560 (1927).
79. Davisson, C. & Germer, L. Diffraction of electrons by a single crystal of nickel. *Science* **30,** 705–741 (1927).
80. Davisson, C. & Germer, L. Reflection of electrons by a crystal of nickel. *PNAS* **14,** 317–322 (1928).
81. *Physics 1922-1941: Nobel lectures including presentations, speeches and laureates' biographies* (Elsevier Publishing Company, Amsterdam, 1965).
82. Davisson, C. Are electrons waves? *Journal of the Franklin Institute* **206,** 597–623 (1928).
83. Hecht, E. *Optics* 5th ed. (Pearson, 2017).
84. Brown, D. Close-range camera calibration. *Photogrammetric engineering* **37,** 855–866 (1971).
85. Zhang, Z. *A Flexible New Technique for Camera Calibration* Technical Report MSR-TR-98-71 (Microsoft Research, 1998).
86. Brown, D. *An advanced reduction and calibration for photogrammetric cameras* Report AD431886 (Air force Cambridge research laboratories. Office of aerospace research, Bedford, Massachusetts, 1964).
87. Snyder, J. *Map projections: A working manual* Report 1395 (U.S. Government Printing Office, Washington, D.C., 1987).
88. Bechtel, J., Smith, W. L. & Bloembergen, N. Two-photon photoemission from metals induced by picosecond laser pulses. *Phys. Rev. B* **15,** 4557–4562 (1977).
89. Petek, H. & Ogawa, S. Femtosecond time-resolved two-photon photoemission studies of electron dynamics in metals. *Progress in Surface Science* **56,** 239–310 (1997).
90. Bauer, M. & Aeschlimann, M. Dynamics of excited electrons in metals, thin films and nanostructures. *Journal of Electron Spectroscopy and Related Phenomena* **124,** 225–243 (2002).
91. Zhu, X. Electron transfer at molecule-metal interfaces: a two-photon photoemission study. *Annu. Rev. Phys. Chem.* **53,** 221–47 (2002).
92. Zhu, X. Electronic structure and eletron dynamics at molecule-metal interfaces: implications for molecule-based electronics. *Surface Science Reports* **56,** 1–83 (2004).
93. Ueba, H. & Gumhalter, B. Theory of two-photon photoemission spectroscopy of surfaces. *Progress in Surface Science* **82,** 193–223 (2007).
94. Moulton, P. Spectroscopic and laser characteristics of Ti:Al_2O_3. *J. Opt. Soc. Am. B* **3,** 125–133 (1986).
95. Strickland, D. & Mourou, G. Compression of amplified chirped optical pulses. *Optics Communications* **56,** 219–221 (1985).
96. Akhmanov, S., Kovrigin, A., Piskarskas, A., Fadeev, V. & Khokhlov, R. Observation of parametric amplification in the optical range. *JETP Lett.* **2** (1965).
97. Franken, P., Hill, A., Peters, C. & Weinreich, G. Generation of optical harmonics. *Phys. Rev. Lett.* **7** (1961).
98. Cerullo, G., Nisoli, M. & Silvestri, S. *Generation of 11-fs pulses tunable across the visible by optical parametric amplification* in *CLEO'98* (1998), 18–19.
99. R.Fork, Martinez, O. & Gordon, J. Negative dispersion control using pairs of prisms. *Optics Letters* **9,** 150–152 (1984).
100. *Application note. Prism compressor for ultrashort laser pulses* Newport (2006).

101. Ferray, M., L'Huillier, A., Li, X. F., Lompre, L. A., Mainfray, G. & Manus, C. Multiple-harmonic conversion of 1064 nm radiation in rare gases. *J. Phys. B: At. Mol. Opt. Phys.* **21,** L31–L35 (1988).
102. Lewenstein, M., Balcou, P., Ivanov, M. Y., L'Huillier, A. & Corkum, P. B. Theory of high-harmonic generation by low-frequency laser fields. *Phys. Rev. A* **49,** 2117–2132 (1994).
103. Nisoli, M., Silvestri, S. D. & Svelto, O. Generation of high energy 10 fs pulses by a new pulse compression technique. *Applied Physics Letters* **68,** 2793–2795 (1996).
104. Frassetto, F., Cacho, C., Froud, C., Turcu, I. E., Villoresi, P., Bryan, W., Springate, E. & Poletto, L. Single-grating monochromator for extreme-ultraviolet ultrashort pulses. *Optics Express* **19,** 19169–19181 (2011).
105. Thibault, P. & Rankenburg, I. Optical diffraction microscopy in a teaching laboratory. *American Journal of Physics* **75,** 827–832 (2007).
106. Devebec, P. & Malik, J. *Recovering high dynamic range radiance maps from photographs* in *Proc. 24th Annual Conf. Computer Graphics and Interactive Techniques* (ACM, New York, 1997).
107. Zuerch, M., Kern, C. & Spielmann, C. XUV coherent diffraction imaging in reflection geometry with low numerical aperture. *Optics Express* **21,** 21131–47 (2013).
108. Patt, M., Wiemann, C., Weber, N., Escher, M., Gloskovskii, A., Drube, W., Merkel, M. & Schneider, C. M. Bulk sensitive hard x-ray photoemission electron microscopy. *Review of Scientific Instruments* **85,** 113704 (2014).
109. Ules, T., Lüftner, D., Reinisch, E. M., Koller, G., Puschnig, P. & Ramsey, M. G. Orbital tomography of hybridized and dispersing molecular overlayers. *Phys. Rev. B* **90** (2014).
110. Lüftner, D., Milko, M., Huppmann, S., Scholz, M., Ngyuen, N., Wiessner, M., Schöll, A., Reinert, F. & Puschnig, P. CuPc/Au(110): Determination of the azimuthal alignment by a combination of angle-resolved photoemission and density functional theory. *Journal of Electron Spectroscopy and Related Phenomena* **195,** 293–300 (2014).
111. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G. & Suetens, P. Multimodality Image Registration by Maximization of Mutual Information. *IEEE Transactions on Medical Imaging* **16,** 187–198 (1997).
112. Bengtsson, E., Wählby, C. & Lindblad, J. Robust Cell Image Segmentation Methods. *Pattern Recognition and Image Analysis* **14,** 157–167 (2004).
113. Graf, M. *et al.* The impact of metalation on adsorption geometry, electronic level alignment and UV-stability of organic macrocycles on TiO2(110). *Nanoscale* **9** (2017).
114. Tian, Q. & Huhns, M. N. Algorithms for Subpixel Registration. *Computer vision, graphics and image processing* **35,** 220–233 (1986).
115. Brown, L. A survey of image registration techniques. *ACM Computing Survey* **24,** 325–376 (1992).
116. Dvornychenko, V. N. Bounds on (Deterministic) Correlation Functions with Application to Registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-5** (1983).
117. Lunt, R. Measurement of the Mean Inner Potentials of Anthracene and Naphtalene. *PRL* **102,** 065504 (2009).
118. Greif, M., Castiglioni, L., Seitsonen, A., Roth, S., Osterwalder, J. & Hengsberger, M. Photoelectron diffraction in the x-ray and ultraviolet regime: Sn-phthalocyanine on Ag(111). *Phys. Rev. B* **87,** 085429 (2013).
119. Dauth, M., Graus, M., Schelter, I., Wiessner, M., Schöll, A., Reinert, F. & Kümmel, S. Perpendicular Emission, Dichroism, and Energy Dependence in Angle-Resolved Photoemission: The Importance of The Final State. *Physical Review Letters* **117,** 183001 (2016).
120. Krüger, P. Photoelectron Diffraction from Valence States of Oriented Molecules. *Journal of the Physical Society of Japan* **87,** 061007 (2018).
121. Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P. & Marchal, G. *Automated multimodality image registration based on information theory* in *Proc. 14th Int. Conf. Information Processing in Medical Imaging; Computational Imaging and Vision 3* (Kluwer, 1995), 263–274.
122. Studholme, C., Hill, D. L. G. & Hawkes, D. J. *Automated 3D Registration of Truncated MR and CT Images of the Head* in *Proceedings of the British Machine Conference* (BMVA Press, 1995), 3.1–3.10.

123. Wells III, W., Viola, P., Atsumi, H., Nakajima, S. & Kikinis, R. Multi-modal volume registration by maximization of mutual information. *Med. Image Anal.* **1,** 35–51 (1996).

124. Viola, P. & III, W. M. W. Alignment by Maximization of Mutual Information. *International Journal of Computer Vision* **24,** 137–154 (1997).

125. Mattes, D., Haynor, D. R., Vessele, H., Lewellen, T. K. & Eubank, W. PET-CT image registration in the chest using free-form deformations. *IEEE Transactions on Medical Imaging* **22,** 120–128 (2003).

126. Woods, R., Mazziotta, J. & Cherry, S. MRI-PET Registration with Automated Algorithm. *Journal of Computer Assisted Tomography* **17,** 536–546 (1993).

127. Cover, T. & Thomas, J. *Elements of information theory* (John Wiley and Sons, Inc., New York, 1991).

128. Unser, M., Aldroubi, A. & Eden, M. Fast B-spline transforms for continuous image representation and interpolation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13,** 277–285 (1991).

129. Thevenaz, P. & Unser, M. *Spline pyramids for inter-modal image registration using mutual information* in *Wavelet applications in signal and image processing* (eds Aldroubli, A., Laine, A. F. & Unser, M.) **3169** (1997), 236–247.

130. Legg, P. A., Rosin, P. L., Marshall, D. & Morgan, J. E. Improving accuracy and efficiency of mutual information for multi-modal retinal image registration using adaptive probability density estimation. *Computerized Medical Imaging and Graphics* **37,** 597–606 (2013).

131. Scott, D. On optimal and data-based histograms. *Biometrika* **66,** 605–610 (1979).

132. Scott, D. *Multivariate density estimation: theory, practice and visualization* (John Wiley and Sons, Inc., 1992).

133. Styner, M., Brechbühler, C., Székely, G. & Gerig, G. Parametric Estimate of Intensity Inhomogeneities Applied to MRI. *IEEE Transactions on Medical Imaging* **19,** 153–165 (2000).

134. Sahoo, P. K., Soltani, S. & Wong, A. K. C. A Survey of Thresholding Techniques. *Computer vision, graphics and image processing* **41,** 233–260 (1988).

135. Davis, L. S. A Survey of Edge Detection Techniques. *Computer vision, graphics and image processing* **4,** 248–270 (1975).

136. Zucker, S. W. Region Growing: Childhood and Adolescence. *Computer vision, graphics and image processing* **5,** 382–399 (1976).

137. Beucher, S. & Lantuejoul, C. *Use of watersheds in contour detection* in *International Workshop on image procesing: Real-time Edge and Motion detection and estimation* (Rennes, France, 1979), 2.1–2.12.

138. Meyer, F. & Beucher, S. Morphological segmentation. *Journal of Visual Communication and Image Representation* **1,** 21–46 (1990).

139. Beucher, S. *Segmentation tools in mathematical morphology* in *Image Algebra and Morphological Image Processing* **1350** (1990).

140. Marr, D. & Hildreth, E. Theory of edge detection. *Proc. R. Soc. Lond. B* **207,** 187–217 (1980).

141. G. E. Sotak, J. & Boyer, K. L. The Laplacian-of-Gaussian Kernel: A Formal Analysis and Design Procedure for Fast, Accurate Convolution and Full-Frame Output. *Computer vision, graphics and image processing* **48,** 147–189 (1989).

142. Byun, J., Verardo, M. R., Sumengen, B., Lewis, G. P., Manjunath, B. S. & Fisher, S. K. Automated tool for the detection of cell nuclei in digital microscopic images: Application to retinal images. *Molecular Vision* **12,** 949–60 (2006).

143. Lindeberg, T. Feature Detection with Automatic Scale Selection. *International Journal of Computer Vision* **30,** 79–116 (1998).

144. Al-Kofahi, Y., Lassoued, W., Lee, W. & Roysam, B. Improved Automated Detection and Segmentation of Cell Nuclei in Histopathology Images. *IEEE Transactions on Biomedical Engineering* **57,** 841–842 (2010).

145. Kong, H., Akakin, H. C. & Sarma, S. E. A Generalized Laplacian of Gaussian Filter for Blob Detection and Its Applications. *IEEE Transactions on Cybernetics* **43,** 1719–1733 (2013).

146. Loukas, C. G., Wilson, G. D., Vojnovix, B. & Linney, A. An Image Analysis-Based Approach for Automated Counting of Cancer Cell Nuclei in Tissue Sections. *Cytometry Part A* **55A,** 30–42 (2003).

147. Medyukhina, A., Meyer, T., Heuke, S., Vogler, N., Dietzek, B. & Popp, J. Automated seeding-based nuclei segmentation in nonlinear optical microscopy. *Applied Optics* **52,** 6979–6994 (2013).

148. Maurer, C. R., Qi, R. & Raghavan, V. A Linear Time Algorithm for Computing Exact Euclidean Distance Transforms of Binary Images in Arbitrary Dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25,** 265–270 (2003).

149. Meyer, F. Topographic distance and watershed lines. *Signal Processing* **38,** 113–125 (1994).

150. Sobel, I. & Feldman, G. *A 3x3 Isotropic Gradient Operator for Image Processing* Presented at the Stanford Artifical Intelligence Project (SAIL). 1968.

151. Danielsson, P.-E. & Seger, O. in (ed Freeman, H.) (Elsevier, 2012).

152. Wählby, C., Sintorn, I.-M., Erlandsson, F., Borgefors, G. & Bengtsson, E. Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections. *Journal of Microscopy* **215,** 67–76 (2004).

153. Frisch, M. J. *et al. Gaussian 09, Revision B.01* Gaussian Inc. Wallingford CT. 2010.

154. Baumberger, F., Hengsberger, M., Muntwiler, M., Shi, M., Krempasky, J., Patthey, L., Osterwalder, J. & Greber, T. Localization of surface states in distorted step lattices. *Phys. Rev. Lett.* **92,** 196805 (2004).

155. Fadley, C. Photoelectron diffraction: space, time, and spin dependence of surface structures. *Surface Review and Letters* **4,** 421–440 (1997).

156. Hutter, J., Iannuzzi, M., Schiffmann, F. & VandeVondele, J. Atomistic simulations of condensed matter systems. *Wiley Interdiscip. Rev.: Comp. Mol. Sci.* **4,** 15–25 (2014).

157. Goedecker, S., Teter, M. & Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B: Condens. Matter Mater. Phys.* **54,** 1703–1710 (1996).

158. VandeVondele, J. & Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **127,** 114105–114113 (2007).

159. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77,** 3865–3868 (1996).

160. Sabatini, R., Gorni, T. & de Gironcoli, D. Nonlocal van der Waals density functional made simple and efficient. *Phys. Rev. B: Condens. Matter Mater. Phys.* **87,** 41108–41111 (2013).

161. Dunning, T. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **90,** 1007–1023 (1989).

162. Perdew, J. P., Burke, K. & Ernzerhof, M. Errata: Generalized gradient approximation made simple. *Phys. Rev. Lett.* **78** (1997).

163. Hyvaerinen, A., Karhunen, J. & Oja, E. *Independent component analysis Adaptive and Cognitive Dynamic Systems: Signal Procvessing, Learning, Communications and Control* (Wiley, 2001).

164. Meier, M. *Time-resolved multi-photon-photoemission on Gallium Arsenide* Bachelor Thesis. University of Zurich (2017).

165. Dougherty, D., Jin, W., Cullen, W., Dutton, G., Reutt-Robey, J. & Robey, S. Local transport gap in C60 nanochains on a pentcene template. *Phys. Rev. B* **77,** 073414 (2008).

166. Amy, F., Chan, C. & Kahn, A. Polarization at the gold/pentacene interface. *Organic electronics* **6,** 85–91 (2005).

167. Repp, J. & Meyer, G. Molecules on insulating films: scanning-tunneling microscopy imaging of individual molecular orbitals. *Phys. Rev. Lett.* **94,** 026803 (2005).

168. Gilton, T., Cowin, J., Kubaik, G. & Hamza, A. Intense surface photoemission: space charge effects and self-acceleration. *J. Appl. Phys.* **68,** 4802 (1990).

169. Schoenhense, B. Multidimensional photoemission spectroscopy - the space-charge limit. *New J. Phys.* **20** (2018).

ACKNOWLEDGMENTS

I would like to thank my supervisor Luca Castiglioni for giving me an opportunity to conduct this exciting research at the University of Zurich, for his constant support, answers to my numerous questions, readiness to help and optimism. It was a great time working together and I did enjoy it! Likewise, I would like to thank Jürg Osterwalder for his wise professional guidance, valuable recommendations and patience. Also, it was a great pleasure to work as a teaching assistant and I am very grateful to have had such an important experience. My work in the laser lab would not be so exciting without Matthias Hengsberger, whom I thank a lot for his support, numerous ideas and especially for his help and patience when I had troubles with lasers. For fruitful discussions about photoemission from molecular adsorbates, their preparation and characterisation, I am grateful to Achim Schöll, who gave numerous unbeatable advices and also generously provided experimental data for validation of phase retrieval algorithms. Also, I thank Hans-Werner Fink for his interest in my project and Peter Puschnig for reviewing this dissertation.

My special thanks are addressed to Tatiana Latychevskaia for introduction into intricacies of phase retrieval, her constant support, openness to my questions and numerous useful advices. Photoemission data used in this work were mostly acquired at the NanoESCA beamline at the Elettra synchrotron in Trieste. Acquiring these data there would not be possible without immense help of Vitaly Feyer, Giovanni Zamborlini and Matteo Jugovac. In turn, interpreting the experimental data would be difficult without DFT calculations, for which I am very grateful to Yeliz Gurdal and Marcella Iannuzzi. Christian Metzger and Manuel Grimm from the University of Würzburg are thanked for interesting discussions and exciting hours spend together in the laser lab and during the beamtime in England. I also thank Christian for his IAC software which he readily consented to provide for the use in this dissertation. Also, I would like to thank our system administrator Roland Bernet for answering my never-ending questions about the institute's linux cluster. Yargo Bonetti from the ETH First Lab is thanked for the design and production of the photolithographic mask used in optical analogue experiment and Andres Kaech for providing an access to the scanning electron microscope. The support of Matthias Muntwiler during my PEARL beamtime is also greatly appreciated. Cephise Cacho and Emma Springate are acknowledged for giving an opportunity to conduct our experiments at the Central Laser Facility and Phil Rice for supporting us there technically.

I also would like to thank all members of the Surface Physics group for nice atmosphere they created around. In particular, my office mates Kay Waltar, Aram Kostanyan, Roland Stania, Rasmus Westerström and Carlo Bernard. Adrian Schuler is thanked for the introduction to our laser system, while Michl Greif is acknowledged for teaching me how to work with Walküre during my first days in the group. Technical help of Thomas Kälin is also greatly appreciated. Karin von Arnx, Matthias Meier, Daniel Schachtler and Patrick Kretz deserve my special thank. It was fun working with you in the lab and I do appreciate your help during the beamtimes. Thomas Greber is thanked for his special ability to ask right questions, sense of humour and very helpful advices. And all administrative questions would not be solved so quickly and easily without kind help of Carmelina Genovese, Monika Röllin, Ruth Halter and Brigitte Freund, whom I am very grateful for helping me at different times on different occasions.

Last but not least, I would like to thank my friends Anya Balbekova, Sonya Blokhina, Dasha Davydova, Artem Ivanov, Slava Kharitonov, Roma Kiselev, Anya Medyukhina, Dima Michel, Anton Sergeyev and Dima Sivun. Thanks for joyful times spend together, for having fun while skiing and traveling! Also, I thank Roma Kiselev for proof-reading my dissertation. Olya Bespalova, Serezha Bespalov, Alsu Gazizulina, Rizalina Mingazheva, Lyosha Nekrasov, Masha Stamatova, Masha Vostrikova and many others from *Zürcher Assoziation der Russischsprachigen Studierenden* are also thanked for happy times spend together! A very special thank you goes to my class-mates Tina Sierova and Kolya Alyoshyn for being my inspiring friends for almost more than 20 years! Another special thank you is reserved for Ulla and Ruedi Wolfensberger and for all Wolfis. You are incredible people and I am so happy to have met all of you! My warmest thanks are addressed to my teachers, in particular, Elena Nikolayevna Lyagushenko and Alexandr Georgievich Nerukh for my love to mathematics, and Evgeniy Nikolaevich Odarenko and Eduard Ivanovich Chernyakov for my passion to optics and solid state physics. Finally, I am immensely grateful to my dear parents for everything they did for me in my life. I am who I am today because of your love, care and believe. Thank you!!!

CIRRICULUM VITAE

PERSONAL

Name	Pavlo KLIUIEV
Date of Birth	23 May 1990
Place of Birth	Lugansk, Ukraine
Citizenship	Ukraine

EDUCATION

2014–2018 PhD in Physics · University of Zurich · Switzerland

- PhD Thesis "Reconstruction of Molecular Orbitals from Photoemission Data with Iterative Phase Retrieval Algorithms". Supervisors: Dr. Luca Castiglioni and Prof. Dr. Jürg Osterwalder

2011–2014 Master of Science in Photonics · Friedrich-Schiller-University of Jena · Germany

- GPA: 1.5/5.0 (1.0 is the highest) · Master Thesis "Gap modes-coupling of plasmonic nanoparticles with a gold film". Supervisors: PD Dr. Wolfgang Fritzsche and Prof. Dr. Thomas Pertsch

2007–2011 Bachelor of Optical Engineering · Kharkiv National University of Radioelectronics, Ukraine

- Honours Degree · Bachelor Thesis "Impact of non-radiating recombination parameters on current-voltage characteristics of bulk heterojunction solar cells". Supervisor: Dr. Alexey V. Shulika

1997–2007 Secondary Education · Gymnasium No. 42 · Lugansk, Ukraine

- GPA: 11.6/12.0 (12.0 is the highest) · Specialisation in Mathematics · Honours and Golden Medal

PROFESSIONAL EXPERIENCE

2014–2018 Research and Teaching Assistant · University of Zurich · Switzerland

- Research on molecular orbital tomography
- Supervision of Master and Bachelor students, teaching assistance in practical exercises on general and condensed matter physics.

2012–2014 Research Assistant · Leibniz-Institute of Photonic Technology · Jena, Germany

- Imaging of gold nanoparticles with dark field optical microscopy
- Modelling of plasmonic nanostructures in COMSOL Multiphysics

2013 Visiting Master student · University of Jyväskylä · Jyväskylä, Finland

- Imaging of gold nanoparticles with near field optical microscopy

2012 Intern · Institute of Applied Physics · Jena, Germany

- Design and construction of a setup for manipulation of gold nanoparticles within a photonic crystal cavity

LANGUAGES

English *German*	Fluent
Russian *Ukrainian*	Native speaker